Impact of plant domestication on spermosphere and rhizosphere microbiome composition

Juan E. Pérez-Jaramillo

Juan E. Pérez-Jaramillo

Impact of plant domestication on spermosphere and rhizosphere microbiome composition

The study described in this thesis was performed at the Netherlands Institute of Ecology, NIOO-KNAW –Wageningen –The Netherlands; practical work was also performed at Universidad de Antioquia, Medellín, Colombia, and the Brazilian Agriculture Research Corporation, Embrapa Meio Ambiente, Jaguariúna, São Paulo State, Brazil.

Design of the cover: Loes Kema

Printed by GVO drukkers & vormgevers B.V. ||www.gvo.nl

ISBN: 978-94-6332-475-5

Impact of plant domestication on spermosphere and rhizosphere microbiome composition

Proefschrift
ter verkrijging van
de graad van Doctor aan de Universiteit Leiden
op gezag van Rector Magnificus Prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 28 maart 2019
klokke 15:00 uur

door

Juan E Pérez-Jaramillo
geboren in 1984 in Jericó, Colombia

PROMOTIECOMMISSIE

Promotor:	Prof. Dr. J.M. Raaijmakers
Co-promotor:	Dr. V.J. Carrión
Overige leden:	Prof. Dr G.P. van Wezel
	Prof. Dr J.A. van Veen
	Prof. Dr J. Falcao-Sales
	Prof. Dr G. Kowalchuk
	Dr D. Bulgarelli

This research was conducted under the auspices of the Colombian Department of Science, Technology and Innovation – COLCIENCIAS.

Table of Contents

Chapter 1

General introduction and thesis outline

A current paradigm in plant science is that plant fitness partly relies on the association with the root- and shoot-associated microbiota (Vandenkoornhuyse *et al*., 2015; Martin *et al*., 2017; Wallenstein, 2017). Across the different stages of plant development, an intricate chemical communication occurs between plants and microbes, resulting in the selection, enrichment and activation of specific microbial groups (Badri *et al*., 2009; Aleklett and Hart, 2013; Lebeis *et al*., 2015). The plant's influence on the soil microbiota starts immediately after seed imbibition, when organic and inorganic compounds are released by the germinating seed in its direct surroundings, a compartment referred to as the spermosphere (Lemanceau *et al*., 2017; Nelson, 2018) After radicle protrusion and root development, plants exert a strong influence on the surrounding soil promoting an overall change in its physical, chemical and biological nature (Badri and Vivanco, 2009; Jones *et al*., 2009). This root-soil interface, referred to as the rhizosphere, is a dynamic, complex and multifaceted environment that is critical for plant growth and health (McNear Jr., 2013). The rhizosphere is home to a diverse group of (micro)organisms, which includes viruses, bacteria, archaea, protozoa, organotrophic and mycorrhizal fungi (Buée *et al*., 2009; Thompson *et al*., 2017). The taxonomically and functionally diverse microbial communities inhabiting and interacting in the rhizosphere are referred to as the rhizosphere microbiome (Boon *et al*., 2013). The complex interplay in the rhizosphere microbiome can result in neutral, beneficial or deleterious effects on plant growth. For instance, inhibition of soil-borne pathogens by specific beneficial members of the rhizosphere microbiome is of pivotal importance for plant health and productivity (Raaijmakers *et al*., 2009). Beneficial effects provided to the plant by specific members of the rhizosphere microbiome can also be attributed to nutrient acquisition, hormone-mediated plant growth promotion, induced systemic resistance and tolerance to abiotic stresses (Mendes *et al*., 2013).

Soil and plant influence rhizosphere microbiome assembly

Soil physical-chemical characteristics and plant genotypic traits are the main drivers of microbiome assembly in the rhizosphere (Garbeva *et al*., 2004; Berg and Smalla, 2009; Delgado-Baquerizo *et al*., 2016). As soil is an enormous reservoir of microbial diversity and serves as the inoculum for rhizosphere microbiome assembly, abiotic characteristics of the soil environment, such as pH, organic carbon content, moisture availability, texture and structure, shape the composition of the rhizosphere microbiome (Fierer, 2017). Also the plant itself shapes the rhizosphere microbiome via the release of root exudates, volatile compounds, mucilage, and border cells into the surrounding soil (Jones *et al*., 2009). Plant species modulate rhizosphere assembly in a genotype-dependent way (Ofek *et al*., 2014; Ofek-Lalzar *et al*., 2014). Interestingly, phylogenetically closely related plant species often present more similar rhizosphere microbiome compositions. For instance, in a study with maize inbred lines, teosinte (*Zea mays* ssp. *parviglumis*), sorghum and wheat, Bouffaud *et al*. (2014) showed that the composition of the microbiome was more divergent with increasing phylogenetic distances between the plant species tested. The intensity of this plant genotype effect can be stronger for specific plant species on the same soil type as was shown by Turner *et al*. (2013b) by RNA-based metatranscriptomics of the rhizosphere microbiomes of pea, oat and wheat. How plant species and cultivars of a given species affect the rhizosphere microbiome composition and activities is still largely unknown. Nevertheless, there are several studies that give some first insight into the genetic and chemical mechanisms underlying rhizosphere microbiome assembly. For instance, it was found in maize that two plant-genes (*su1*, *sh2*), involved in different carbon storage patterns affecting root exudate composition, promoted different microbial communities in the rhizosphere (Aira *et al*., 2010). Differential recruitment of beneficial microorganisms was also shown for wheat, where the root colonization by specific antifungal *Pseudomonas* species was highly cultivar dependent (Mazzola *et al*., 2004;

Kwak *et al*., 2012). Another very good example of host genotype-dependent recruitment was shown for eleven *Arabidopsis thaliana* accessions, each with a distinct rhizosphere community composition, probably as a result of genotype-specific patterns of root exudation (Micallef *et al*., 2009b).

Plant exudates and microbiome assembly

Seed and root exudates comprise low-molecular weight compounds, i.e. sugars, amino acids, organic acids, phenolics and several secondary metabolites, but also high-molecular weight compounds like proteins and mucilage (Nelson, 2004; Bais *et al*., 2006; Carvalhais *et al*., 2011; Kawasaki *et al*., 2016; Sasse *et al*., 2018). Other compounds commonly found in natural blends of plant exudates are flavonoids, strigolactones and salicylic acid (Cesco *et al*., 2010; Lebeis *et al*., 2015; Zwanenburg *et al*., 2016). Plant genotype, soil type and abiotic factors largely determine the type and quantity of exudates released by seeds and roots (Lareen *et al*., 2016; Mönchgesang *et al*., 2016; Iannucci *et al*., 2017; Lemanceau *et al*., 2017). Through active and passive mechanisms of exudation, plants impact soil physical-chemical conditions (Rohrbacher and Marc St-Arnaud, 2016; Wang *et al*., 2017) and the composition and activities of spermosphere and rhizosphere microbiomes (Pérez-Jaramillo *et al*., 2016; Nelson, 2018). Depending on the type and amount of exudates released, the plant may select and enrich a subset of microbes from the diverse microbial populations inhabiting the soil, or may also act as modulators of soil biogeochemical processes (Hassan and Mathesius, 2012; Coskun *et al*., 2017). Symbiotic associations with mycorrhizal fungi and rhizobia are initially triggered by plant exudates, and followed by an intricate chemical communication (Chagas *et al*., 2018). For instance, under N and P limitations, plants release strigolactones into the rhizosphere, which induce several responses in mycorrhizal fungi such as spore germination, hyphal growth and hyphal branching, triggering the symbiotic association with the plant (Lanfranco *et al*.,

2017; Waters *et al*., 2017). Similarly, in the nodulation process of legumes, the plant root secretes specific flavonoids which are recognized by bacterial Nod factors. This initial step triggers the formation of intracellular structures in the plant that allows compatible rhizobia to enter the inner root tissues, form the nodule and start the N-fixation process (Liu and Murray, 2016). However, not only symbiotic rhizobia inhabit legume nodules. In the model legume *Lotus japonicus,* it was shown that next to Rhizobiales also other bacterial taxa such as Burkholderiales, Flavobacteriales, Pseudomonadales, and Actinobacteridae are found associated with the nodules (Zgadzaj *et al*., 2016). In the same study, it was shown that an impaired symbiosis ability of the host plant can have a large impact on the bacterial community of the root and rhizosphere compartments (Zgadzaj *et al*., 2016). Also plant hormones released to the soil environment can impact on the root microbiome assembly. Specifically, it has been shown that wild type Arabidopsis levels of salicylic acid (SA) promoted a particular set of families in the root compartment while limiting colonization of several other bacterial taxa as compared to Arabidopsis mutants deficient in SA biosynthesis (Lebeis *et al*., 2015). Phenolic acids can also impact the composition of the endophytic microbiome. For instance, cinnamoyl-CoA reductase (*CCR*)-down-regulated poplar trees, which simultaneously display reduced lignin levels and accumulation of extractable phenolic compounds in the xylem, showed an increase in bacterial cell counts as compared to wild type poplar trees. This finding suggested that the capacity to degrade phenolic compounds, in particular pherulic acid, was enhanced in the endosphere of *CCR*-down-regulated poplar trees, directly affecting the endophytic bacterial community (Beckers *et al*., 2016).

Plant exudates may be also released in order to trigger microbiome-dependent defensive responses. Chapelle *et al.* (2016), found an overrepresentation of ppGpp metabolism, which is a mechanism associated with stress perception and stress response regulation in

bacteria, in the rhizosphere microbiome of sugar beet grown in a soil suppressive to the fungal root pathogen *Rhizoctonia solani*. They hypothesized that this defensive response may have been triggered by plant-derived compounds following initial root infection by *R. solani* or directly by organic acids produced by the pathogenic fungus prior to or during infection, such as oxalic and phenylacetic acid (Chapelle *et al*., 2016). Besides soluble compounds, plants can also release volatile organic compounds (VOCs) to communicate with bacterial communities. In a recent study by Schulz-Bohm *et al.* (2018), it was shown that roots of *Carex arenaria* infected with the fungal pathogen *Fusarium culmorum* emitted a blend of VOCs that attracted bacteria with antifungal properties. In conclusion, plants can release several types of compounds into the soil environment to directly impact the composition of the microbiome assembly. Furthermore, these and other studies indicated that plants under attack alter their exudation profiles to recruit specific members of the microbiome with the ability of antagonize subsequent pathogen attacks. For instance, infection of Arabidopsis leaves with *Pseudomonas syringae* pv *tomato* induced an increased secretion of malic acid from the roots, promoting the colonization and biofilm formation by the beneficial rhizobacterium strain *Bacillus subtilis* FB17 (Rudrappa *et al*., 2008). Similarly, under foliar pathogen infection, Arabidopsis plants activated a salicylic acid dependent defensive response in order to recruit specific bacterial taxa in the rhizosphere which provided systemic protection against the pathogen (Berendsen *et al*., 2018). This 'cry for help' was already well-described aboveground as an indirect defense mechanism of plants in response to insect herbivory (Dicke, 2009) but is less well studied belowground. One of the first conclusive studies on the "crying for help" hypothesis belowground was shown for maize. Upon attack by insect larvae, a volatile sesquiterpene (E)-β-caryophyllene was emitted via the roots attracting an entomopathogenic nematode (Rasmann *et al*., 2005). It is highly likely that the same principle holds for plant-pathogen interactions in the rhizosphere environment, where

specific members of the microbiome may act as an indirect plant defense mechanism (Cook *et al*., 1995).

The search for beneficial rhizosphere traits

Because the importance of the microbiome for plant growth and health has been largely underestimated, there is a renewed interest in identifying plant and microbial traits involved in the beneficial relationships between microbes and plants (Busby *et al*., 2017; Kroll *et al*., 2017; Wallenstein, 2017). One of the first attempts to identify plant traits and genes involved in beneficial interactions with soil microorganisms, other than rhizobia and arbuscular mycorrhizal fungi (AMF), was in a study by Smith *et al*. (1997) with six tomato lines and the biocontrol strain *Bacillus cereus* UW85. In this study, the authors found significant differences between the tomato lines for resistance to the pathogen *Pythium torulosum* and for support of root colonization and disease suppression by the biocontrol strain UW85. Subsequent analysis of recombinant inbred line (RIL) populations, derived from a cross between a cultivated tomato and a wild tomato relative, showed that three quantitative trait loci (QTL) were associated with disease suppression by *B. cereus* UW85, and three other QTL were associated with the growth of *B. cereus* on the seed (Smith *et al*., 1999). More recently, genetic loci have been identified for their involvement in the ability of Arabidopsis accessions to respond to the activity of plant growth-promoting bacteria (Wintermans *et al*., 2016). Furthermore, Mendes *et al*. (2018) showed that resistance breeding in common bean against the root pathogen *Fusarium oxysporum* also had unintentionally co-selected for plant traits that support rhizosphere microbes with potential antagonistic activities to this fungal root pathogen. On the other side, some genes and traits related to bacterial adaptation to plants have been also recently characterized. Briefly, gene categories related to carbohydrate metabolism and transport were enriched in the genomes of several plant- and root- associated bacteria, while mobile

genetic elements were found underrepresented in plant-associated taxa (Levy *et al*., 2018). However, several conceptual and experimental efforts are yet to be made to enable the use of these traits in rhizosphere-based plant breeding programs (Bakker *et al*., 2012). These efforts include the identification of the specific plants traits involved and complementing the still limited knowledge of the microbial diversity responsible for the beneficial effects on plant growth and health. Meta-'omics approaches have opened new possibilities to also identify those microbial groups and their functional genes in the rhizosphere involved in biocontrol or in plant growth promotion that were invisible due to poor culturability (Ramírez *et al*., 2017; Thompson *et al*., 2017). These technologies also allow us to unravel changes in the abundance, evenness and richness as well as in their functions, and provide a first step in the selection of prominent plant genotypes that are better equipped in shaping a beneficial rhizosphere microbiome.

Plant domestication and microbiome assembly

Over the past decades, plant breeders have exploited genes from wild relatives of modern crop species to improve plant growth and health (Hajjar and Hodgkin, 2007). Similarly, entomologists explore native habitats and wild ancestors of crop plants to identify natural enemies of insect pests (Chen, 2016). In plant microbiome research, however, relatively few efforts have been made to study the identity and diversity of beneficial microbial communities present in the native habitats of ancestors of modern agricultural crop species. Smith *et al*. (1999) postulated that plant genotypic traits involved in beneficial plant-microbe interactions may have been lost in the process of plant breeding, primarily as a consequence of the targeted selection of a few specific traits related with higher yields or other profitable plant traits. In fact, it is known that loss of genetic diversity is a common feature of domesticated crops as compared to their wild relatives (Bitocchi *et al*., 2013). This reduction in genetic diversity could have undermined the abilities of

modern plant cultivars to interact with and benefit from the microbiome. Genes involved in active root exudation and/or root development could have been inadvertently lost in domesticated plant accessions. Therefore, I postulate that wild relatives of crop plants, which are more genetically diverse, represent a valuable source of genotypic and phenotypic traits associated with the recruitment and assembly of the microbiome. Furthermore, domestication also led to environmental changes associated with agricultural practices which in turn made biotic interactions with belowground diversity presumably less relevant to sustain plant growth and health (Wissuwa *et al*., 2009). Consequently, I postulate that native habitats may harbor a soil microbial diversity that co-evolved with wild relatives of crop plants and that agricultural management and habitat expansion of domesticated varieties caused a decoupling between plants and their microbiome. This could have led to the loss of microbial diversity, the 'missing plant microbes', which we hypothesized were formerly a functionally important component of the microbiome of wild plants in their native habitats. This concept has been explored in human microbiome research, where antibiotic overuse and modern lifestyle also are proposed to have caused a loss of several components of the microbiota (Blaser, 2017). In this context, I postulate that modern crops grown in soils under agricultural management have lost specific members of the rhizosphere microbiome typically found on roots of their wild relatives growing in the native soils in the centres of origin.

Common bean as a 'model' in plant microbiome and domestication research

Common bean (*Phaseolus vulgaris* L.) is currently the most important legume crop, grown worldwide in a broad range of environmental conditions. It is of high importance in developing countries because it is a good source of protein, carbohydrate and micronutrients. Furthermore, common bean is the basis of subsistence economy for low income farmers especially in Latin America and Africa, currently the largest common

bean producing regions (Broughton *et al*., 2003; Akibode and Maredia, 2011; CGIAR, 2013). The common bean rhizosphere microbiome has been studied primarily because of the symbiotic associations with nitrogen-fixing bacteria of the genus *Rhizobium*. Briefly, five *Rhizobium* species have been recognized as microsymbionts of *P. vulgaris* (Amarger, 2001), and the species *R. etli* as the most commonly associated with beans in the Americas (Aguilar *et al*., 2004). The regulation of this symbiotic association is dependent on compounds produced by the plant (flavonoids) for the induction of nodulation genes in the bacterial species, which in turn secrete nodulation factors essential for nodule formation (Schultze and Kondorosi, 1998). Apart from this well-studied plant-microbe association, the relation between rhizobia and several nonsymbiotic bacteria in the rhizosphere of common bean has been assessed, searching for an increment in nodule formation mediated by bacterial inoculants (Burdman *et al*., 1996; Petersen *et al*., 1996). To our knowledge, the differences between wild relatives, landraces and cultivars of common bean in the composition and functions of the microbial communities in the rhizosphere have not been established, nor the relative influence of plant genotype and soil type.

Wild common bean (*Phaseolus vulgaris*) originated in central Mexico (Bitocchi *et al*., 2012; Desiderio *et al*., 2013) and from there spread throughout Central and South America (Gepts, 1998). This wide distribution led to the establishment of two different genetic pools: Mesoamerican and Andean (Gepts and Bliss, 1985). Additional to these genetic pools, a third intermediate genetic pool can be distinguished composed of wild common bean accessions from Ecuador and northern Peru with a combination of both Mesoamerican and Andean alleles (Debouck *et al*., 1993) (Fig. 1). The Mesoamerican and Andean wild common bean populations were the basis for two independent domestication processes, which presumably occurred in the Oaxaca Valley in Mexico for

Mesoamerica and eastern Bolivia or northern Argentina for the Andes (Beebe *et al*., 2001; Bitocchi *et al*., 2013). After domestication, common bean started a journey across the American continent with human groups as its main dispersion force and successfully adapted to several dissimilar environments and human preferences, leading to several morphological and physiological changes (Gepts and Debouck, 1991). Simultaneously, domesticated common bean races experienced a reduction in genetic diversity as compared to its wild counterparts, a fact that encouraged plant breeders to look into wild common bean genotypes for useful traits that could be used for the improvement of domesticated races. The broadening of the cultivable gene pool through the use of wild relatives or through the recombination of Mesoamerican and Andean gene pools has been postulated as a response to abiotic and biotic stresses, and also as a strategy to improve the yield of commercial bean varieties (Gepts, 1998; Singh, 2001).

Fig. 1. Distribution of wild common bean (*Phaseolus vulgaris*) in Central and South America. Accessions of common bean have been collected from Mexico to Argentina and three genetic pools can be recognized based on phenotypic and molecular data: Mesoamerican, Andean and an intermediate gene pool

in northern Peru and Ecuador. Adapted from "Origin and evolution of common Bean: Past events and recent trends", Gepts, (1998).

Wild beans have coevolved with several pests and diseases, symbiotic organisms such as rhizobia and mycorrhizal fungi, and at the same time acquired adaptations to several abiotic conditions (Toro *et al*., 1990). Recently, it was proposed that wild relatives could be reservoirs of genes associated with drought tolerance (Cortés *et al*., 2012). An example of successful wild bean gene mining was the transfer of genes associated with high grain yield from wild beans to cultivars of domesticated common bean (Blair *et al*., 2006). In another study, it was observed that when a wild common bean was used as one of the parents, more plants with a higher yield were found in the segregating populations compared to the domesticated parent (Herrera-Flores and Acosta-Gallegos, 2008). However, in a recent survey in which 19 important crops were evaluated with regard to the contributions of wild relatives to their overall crop performance, it was noticed that common bean is amongst those species, together with soybean, pigeon pea, sorghum, lentil, and cowpea, that have not incorporated wild genes into the cultivated varieties, whilst for the remaining 13 crops over 100 traits have been acquired from over 60 wild species (Hajjar and Hodgkin, 2007). The lack of wild traits in modern common bean varieties indicates that from the plant breeding perspective more efforts are yet to be made in order to identify and integrate beneficial traits from wild common bean genotypes into modern cultivars. As mentioned above, these traits could be used to address problems associated with biotic and abiotic stress factors and to improve plant productivity, but also the potential of a still unknown genetic diversity could be used to overcome the future production constraints in the scenery of climate change (Porch *et al*., 2013). Thousands of accessions of modern common bean cultivars, landraces and wild relatives are held in the Program of Genetic Resources from the International Centre for Tropical Agriculture (CIAT) to perform basic and applied research (Toro, 1990; CIAT-PRG, 2018). With the

aforementioned plant material we can determine the inherent rhizosphere microbial diversity of cultivated varieties, landraces and wild relatives, allowing comparisons which can be used to assess whether there is a core microbiome of the numerous common bean accessions. By contrast, it can be also determined whether wild relatives and landraces, which possess a more diverse genetic background, can promote different rhizosphere microbial communities as compared to those present in domesticated modern varieties. Finally, the relative importance of the plant genotype in shaping the rhizosphere microbial diversity can be unraveled with common bean as a model plant and if/how microbiome diversity impacts the plant's overall performance.

Thesis Outline

The overall aim of my thesis was to decipher the impact of plant domestication on the spermosphere and rhizosphere microbiome composition of common bean (*Phaseolus vulgaris*). Common bean is currently the most important legume crop, grown worldwide in a broad range of environmental conditions with an outstanding nutritional importance in developing countries. In order to explore the effect of domestication on the plant microbiome, an extensive literature review was undertaken (**Chapter 2).** This chapter also presents the "Back-to-the-Roots" approach, which is a general workflow to investigate the impact of domestication on microbiome composition and functioning and how you can use that knowledge to reinstate beneficial partnerships in modern cultivars by studying the microbiome of wild plants in native soils. **Chapter 3** links genotypic and phenotypic traits of wild and domesticated common bean accessions with their rhizosphere microbiome composition. Genotypic and root morphology analyses of the common bean accessions, followed by a detailed characterization of the rhizosphere microbiome by 16S rDNA amplicon sequencing, allowed a robust statistical analysis to find associations between genotypic and root phenotypic traits of common bean and the

abundance of certain rhizobacterial families. **Chapter 4** addresses how the transition of common bean from a native soil to an agricultural context impacts the composition of the rhizosphere microbiome. Soil sampling was performed in Colombia, one of the centres of diversification of common bean. Subsequently, the native and the agricultural soils collected were used to grow different wild and modern common bean accessions in a pot trial under controlled conditions. An integral assessment was undertaken, including diversity measurements, niche-neutral abundance modelling, core microbiome assessment, differential abundance and co-ocurrence network analyses to search for the 'missing plant microbes', i.e. microbes that were lost during domestication of common bean including the transition from a native to an agricultural soil. **Chapter 5** describes how domestication of plant species other than common bean impacted the root-associated microbiome. To this end, we analyzed publicly available metagenomic data from microbiome studies on different modern crop cultivars and their wild relatives. Following the same pipeline in this 'computational walk on the wild side', we determined differences in root microbiome composition between wild relatives and their domesticated counterparts. We specifically identified the relative abundance of specific taxa within the Bacteroidetes phylum.

Different plant developmental stages may recruit and enrich different microbial groups. The spermosphere, which is a short-lived, dynamic developmental stage, can also have an impact on microbiome assembly. **Chapter 6** evaluates through amplicon sequencing the magnitude of the spermosphere effect on the bacterial diversity and abundance to assess if a plant genotype-dependent influence is already discernible in the spermosphere of common bean. Subsequently, collection, characterization and quantification of seed exudates were performed. In a further experiment, one of the most abundant amino acids (i.e. glutamate) released by common bean seeds was used as inoculant of a soil solution, in order to evaluate if a single seed exudate can simulate the compositional shifts observed

between the spermosphere microbiomes of wild and modern bean accessions. **Chapter 7** brings together the research findings of this thesis, discusses the implications of the conceptual and experimental work and proposes further directions on how to translate the fundamental knowledge on rhizosphere microbiomes of wild relatives and native habitats to modern agriculture.

Chapter 2

Impact of plant domestication on rhizosphere microbiome assembly and functions

Juan E. Pérez-Jaramillo, Rodrigo Mendes, Jos M. Raaijmakers

Plant Molecular Biology 90(6): 635–644 (2016)

https://doi.org/10.1007/s11103-015-0337-7

Abstract

The rhizosphere microbiome is pivotal for plant health and growth, providing defence against pests and diseases, facilitating nutrient acquisition and helping plants to withstand abiotic stresses. Plants can actively recruit members of the soil microbial community for positive feedbacks, but the underlying mechanisms and plant traits that drive microbiome assembly and functions are largely unknown. Domestication of plant species has substantially contributed to human civilization, but also caused a strong decrease in the genetic diversity of modern crop cultivars that may have affected the ability of plants to establish beneficial associations with rhizosphere microbes. Here, we review how plants shape the rhizosphere microbiome and how domestication may have impacted rhizosphere microbiome assembly and functions via habitat expansion and via changes in crop management practices, root exudation, root architecture, and plant litter quality. We also propose a "back to the roots" framework that comprises the exploration of the microbiome of indigenous plants and their native habitats for the identification of plant and microbial traits with the ultimate goal to reinstate beneficial associations that may have been undermined during plant domestication.

Keywords: rhizosphere microbiome, plant domestication, wild relatives, plant-microbe interactions

Introduction

Plants rely on their rhizosphere microbiome for functions and traits related to plant growth, development and health (Berendsen *et al*., 2012; Mendes *et al*., 2013). Members of the rhizosphere microbiome harbour a range of beneficial properties contributing to nutrient acquisition, enhanced stress tolerance, protection against soil borne pathogens and host immune regulation (Berendsen *et al*., 2012; Bakker *et al*. 2013; Mendes *et al*., 2013; Turner *et al*., 2013a; Berg *et al*., 2014; Lakshmanan *et al*., 2014). In this context, Cook *et al*. (1995) postulated that natural selection resulted in only few examples of plant genetic resistance against belowground pathogens and that plant rely, in part, on the natural defence provided by rhizosphere microorganisms. This is the case for natural disease suppressive soils where specific microbial consortia protect the host from infection (Mendes *et al*., 2011). Assuming that plants depend, at least in part, on the rhizosphere microbiome as a product of natural selection, modern cultivars of crop plants may have lost some of the traits needed to recruit host-specific root microbiota as compared to their wild relatives, which are genetically more diverse and adapted to pre-agricultural soils (Wissuwa *et al*., 2009; Bulgarelli *et al*., 2013). Whether the ability of crop plants to recruit beneficial rhizosphere microbes is undermined by plant domestication and plant breeding is not well known to date. In this review, we discuss the potential influence of plant domestication on rhizosphere microbiome assembly and function, focusing on how domestication may have impacted the ability of modern crops to establish beneficial interactions with the rhizosphere microbiome. Finally, we propose a framework for identification and recovery of beneficial plant-microbe interactions to meet the need for a more sustainable and productive agriculture.

Plant domestication: changes and trade-offs

One of the biggest accomplishments in human history has been the domestication of plants, providing a more continuous food supply and promoting the conformation of sedentary agricultural groups (Purugganan and Fuller 2009). The process of plant domestication involves selection, modification and adoption of wild plants species with useful characteristics for human use (Gepts, 2004). The first changes commonly associated with plant domestication were a large seed size, loss of seed dispersal mechanisms, and determinate growth and apical dominance (Gross and Olsen 2010). Other changes comprise the loss of seed dormancy, decrease of bitter substances in edible structures and changes in photoperiod sensitivity (Purugganan and Fuller 2009). Domestication also led to a reduction in genetic diversity of plant cultivars as shown for common bean (Bitocchi *et al*., 2012; Bittochi *et al*., 2013), rice (Ram *et al*., 2007) and wheat (Haudry *et al*., 2007). Genes associated with desirable phenotypes underwent a diversity loss because only the desired alleles were spread in the subsequent progenies, whilst unwanted diversity of the same allele was inadvertently suppressed (Doebley *et al*., 2006). In addition, genomic regions next to the target genes suffered selective sweeps as was shown for the adjacent regions of the *Y1* phytoene synthase gene for endosperm colour in maize (Palaisa *et al*., 2004) and of the *Waxy* granule-bound starch synthase gene for amylose synthesis in rice (Olsen *et al*., 2006). Thus, a possible side effect of plant domestication is the loss of traits neglected during human selection. In a recent review, Chen *et al*. (2015) indicated that the ability of plants to deal with herbivorous insects is undermined in domesticated crops, in part as a consequence of changes in morphological traits and in levels of secondary metabolites, which make domesticated plants a better resource than their wild relatives. Chen *et al*. (2015) further highlighted that domestication led to lower levels of volatile emissions as compared to wild relatives, which in turn may affect the attraction of natural enemies. Whether plant traits needed to

recruit and sustain beneficial microbial populations in the rhizosphere was also negatively impacted remains to be elucidated.

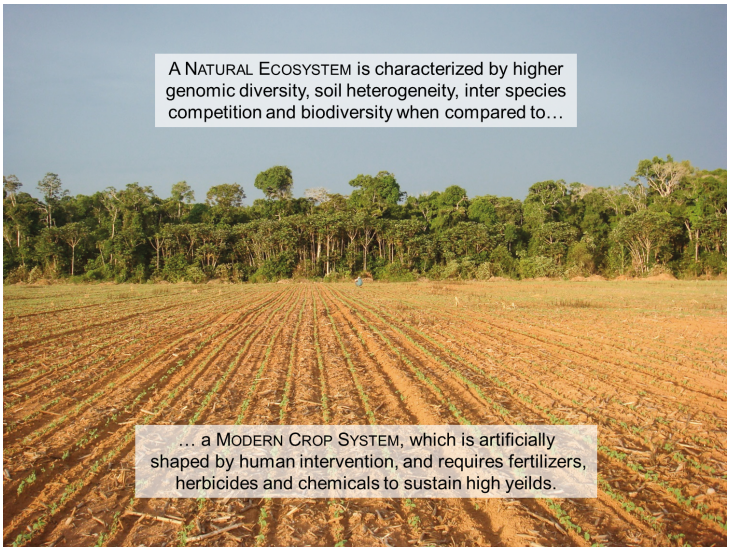


Fig. 1. In this image, the natural ecosystem is illustrated by the native Amazon rainforest (background) that was converted to a modern crop system (front). This conversion leads to changes in the environment and use of management techniques ultimately impacting the rhizosphere microbiome assembly and functions (Photo by L. W. Mendes).

Many of the changes in plant traits during domestication were accompanied by progressive changes in the environment and management practices (Fig. 1). Hence, plant domestication associated with anthropogenic interference to sustain high yields led to low self-support production systems with an increased need for external inputs such as chemical pesticides and fertilizers to overcome problems related to pests and diseases, vulnerability to abiotic stress and nutrient depletion (Matson *et al*., 1997). Moreover, the transition from natural to agricultural systems may have hampered beneficial interactions between plants and microbes due to loss of soil microbial diversity. For instance, it was shown that long-term nitrogen fertilization resulted in the evolution of less-mutualistic rhizobia, providing fewer benefits to the host (Weese *et al*., 2015). Nitrogen amendments

have also been shown to suppress soil respiration and microbial biomass, promoting copiotrophs such as Actinobacteria and Firmicutes while reducing the abundance of oligotrophs such as Acidobacteria and Verrucomicrobia (Ramirez *et al*., 2012). This was substantiated by Rodrigues *et al.* (2013) who showed that conversion of the Amazon rainforest to agriculture resulted in biotic homogenization of soil bacterial communities and reduction of microbial diversity. Fierer *et al*. (2013) further showed that in a native tallgrass prairie ecosystem, bacterial communities did not resemble those harboured by the surrounding cultivated soils where Verrucomicrobia represented more than 50% of the 16S rRNA sequences identified. Also soil attributes can be affected by plant domestication, which in turn influence the soil microbial community composition. García-Palacios *et al*. (2013) demonstrated, in microbial-rich and microbial-poor soils, that plant domestication increased litter quality, resulting in lower C:N ratio and higher NO_3 availability. In addition to changes in the production systems, domesticated lineages experienced range expansions far beyond their centres of origin due to human migrations and trade (Purugganan and Fuller 2009). Hence, the lack of a co-evolutionary trajectory between plants, microbial communities and pathogens in dissimilar agricultural landscapes, made human interventions even more critical to maintain a healthy and productive crop (Fig. 2).

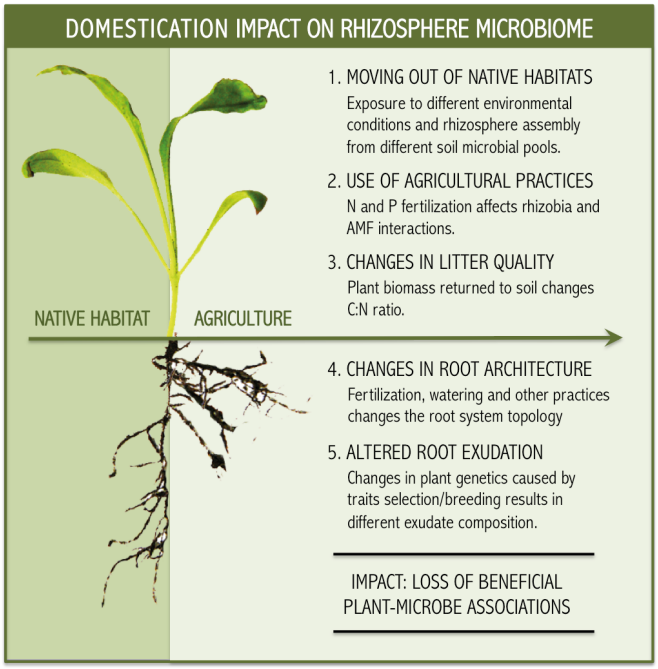


Fig. 2. Changes associated to the domestication process effect plant traits and soil properties undermining rhizosphere microbiome composition and functions.

Effect of plant genotype on rhizosphere microbiome assembly

Plants can modulate their rhizosphere microbiome in a host-dependent way. Each plant species promotes a particular set of rhizosphere microbes (Haichar *et al*., 2008; Turner *et al*., 2013b; Ofek *et al*., 2014). With an increase in the phylogenetic distance between plant species also differences in the composition of their rhizosphere microbial assemblages appear to increase (Wieland *et al*., 2001; Pongsilp *et al*., 2012; Bouffaud *et al*., 2014). Not only different plant species, but also different genotypes of the same species may differ in their rhizosphere microbiome composition. For example, Weinert *et al.* (2011) showed for three different potato cultivars that a portion of the detected OTUs was cultivar-specific and that the *Streptomycetaceae* responded in a cultivar-dependent manner. Similar cultivar-dependent effects were observed for the rhizobacterial communities in the rhizosphere of young potato plants (Inceoglu *et al*., 2011). In a recent

study with 27 modern maize inbred lines, grown in five field environments, Peiffer *et al*. (2013) showed that OTU richness was affected by maize genotypes and that the variation in β-diversity was partially explained by the maize genotype. Similarly, in a study with different barley genotypes, Bulgarelli *et al*. (2015) found that the host genotype accounts for approximately 5.7% of the variance in the rhizosphere microbiome composition. In sweet potato (*Ipomoea batatas*), *Sphingobium*, *Pseudomonas*, *Acinetobacter*, *Stenotrophomonas*, and *Chryseobacterium* were enriched on the low starch genotype as compared to two high starch genotypes (Marques *et al*., 2014). Next to genotype-specific effects, also the plant developmental stage is a strong driver shaping the rhizobacterial community structure. In soybean, *Bradyrhizobium*, *Bacillus* and *Stenotrophomonas* were more abundant at the flowering stage as compared to vegetative and mature stages (Sugiyama *et al*., 2014a). For fungal communities, however, no significant effects of the soybean growth stage were detected (Sugiyama *et al*., 2014b). These effects, however, are not general as some studies highlighted a stronger selective rhizosphere effect at young plant growth stages (Gomes *et al*., 2001; Jin *et al*., 2009; Micallef *et al*., 2009a; Xu *et al*., 2009; Chaparro *et al*., 2014) whilst other studies documented stronger selective effects on the microbiome at flowering (Smalla *et al*., 2001; Inceoglu *et al*., 2010).

Plant root exudates and the recruitment of beneficial microbes

Plants actively release exudates, volatile compounds, border cells and polymers into the soil, a process referred to as rhizodeposition (Jones *et al*., 2009). Root exudates are composed of low-molecular weight compounds, i.e. sugars, amino acids, organic acids, phenolics, secondary metabolites, and high-molecular weight compounds like proteins and mucilage (Badri and Vivanco 2009). For more details on the chemical diversity of compounds in the rhizosphere we refer to other reviews (Bais *et al*., 2006; Badri *et al*., 2009; Moe 2013; Weston and Mathesius 2013; Baetz and Martinoia 2014; Haichar *et al*.,

2014). Root exudates may impact the functioning of the microbial community. For instance, in soil amended with maize mucilage a higher production of N_2O was recorded than in non-amended soil (Mounier *et al*., 2004). Likewise, additions of artificial root exudates (ARE) to a soil microcosm, mimicking maize exudates, promoted nitrate reduction and denitrification (Henry *et al*., 2008). A study with eight *Arabidopsis thaliana* accessions revealed that each accession released a particular set of exudate compounds and that each accession had a distinct rhizobacterial community composition based on RISA and 16S-TRFLP analyses (Micallef *et al*., 2009b). Some root exudates may impact the microbial community structure to a greater extent than other compounds as was shown for organic acids with a 10-22 fold increase in the bacterial taxa while sugars showed only a 2.5 fold increase (Shi *et al*., 2011). When Arabidopsis exudates collected from 18-21 days old plants were applied to a fallow soil, phenolic compounds had a significant positive correlation with the highest number of bacterial OTUs (742) whereas lower number of OTUs were found for amino acids (319), sugar alcohols (166), and sugars (161) (Badri *et al*., 2013). Root exudates such as flavonoids or strigolactones play key roles in symbiotic relationships between plants and rhizobia, mycorrhiza and also parasites (Jones *et al*., 2004; Bouwmeester *et al*., 2007; Bednarek *et al*., 2010; Wang *et al*., 2012; Haichar *et al*., 2014). Likewise, root exudates may impact specific groups of beneficial bacteria referred as plant growth promoting rhizobacteria (PGPR). For example, *Pseudomonas fluorescens* WCS365, a rhizosphere competent colonizer of tomato roots, was chemotactically attracted towards malic acid and citric acid exuded by tomato roots (de Weert *et al*., 2002). Also beneficial rhizobacterium *Bacillus subtilis* FB17 exhibited a positive chemotactic response towards L-malic acid. Interestingly, infection of Arabidopsis leaves with *Pseudomonas syringae* pv *tomato* induced an increased secretion of malic acid from the roots, promoting the colonization and biofilm formation by strain FB17 (Rudrappa *et al*., 2008). Furthermore, foliar pathogens or foliar

treatment with microbe-associated molecular patterns (MAMPs) of Arabidopsis leaves promoted the expression of the root malic acid (MA) transporter (*ALMT1*), stimulating the colonization by *Bacillus subtilis* strain FB17 (Lakshmanan *et al*., 2012). Malic acid and citric acid exuded by watermelon roots were shown to induce motility and root colonization by the PGPR *Paenibacillus polymyxa* SQR-21 (Ling *et al*., 2011). Similar effects of malic acid and citric acid were found for chemotaxis and biofilm formation by *Bacillus amyloliquefaciens* SQR9 in cucumber and for fumaric acid in promoting colonization of banana roots by *Bacillus subtilis* N11 (Zhang *et al*., 2014). Also other compounds found in root exudates may recruit beneficial bacteria. The aromatic compound 2,4-dihydroxy-7-methoxy-2H-1,4-benzoxazin-3(4H)-one (DIMBOA) exuded by maize roots, showed a chemoattractant effect on and an increased root colonization by *Pseudomonas putida* KT2440 (Neal *et al*., 2012). Finally, plant derived compounds may also have an effect on the expression of bacterial antifungal biosynthetic genes. For instance, the expression of *phlA* and *pltA* genes in *Pseudomonas fluorescens* CHA0, involved in the biosynthesis of the antifungal compounds 2,4-diacetylphloroglucinol (DAPG) and pyoluteorin (PLT) respectively, was induced or repressed by 40 different plant-derived compounds, including several plant phenolics and pectin (de Werra *et al*., 2011). Similarly, phenolic and organic acids exuded by barley plants infected with the fungus *Pythium ultimum* induced the expression of the *phlA* gene of *P. fluorescens* CHA0, presumably as a plant systemic response to deal with the pathogen (Jousset *et al*., 2011). Also *Zea mays* subsp. *parviglumis* and European maize lines emitted the volatile sesquiterpene (*E*)-β-caryophyllene via the roots attracting an entomopathogenic nematode in response to insect attack; North American lines failed to release this compound probably as a consequence of the breeding process (Rasmann *et al*., 2005; Köllner *et al*., 2008).

These results exemplify the potential of plants to recruit and activate, via specific components in root exudates, beneficial members of the rhizosphere microbiome. However, plant exudates may also exert a negative effect on belowground communities. In a study evaluating the effect of the invasive weed *Centaurea maculosa* on the composition of arbuscular mycorrhizal fungal (AMF) communities, the abundance and diversity of AMF was reduced compared to native grassland samples (Mummey and Rillig 2006). The same deleterious effect of *C. maculosa* was also shown for the overall soil fungal community. Broz *et al.* (2007) further observed that high density weed populations had a reduced fungal biomass and diversity as compared to low density weed populations mixed with native species. Badri and Vivanco (2009) suggested that root exudates released by invasive weeds disrupt the indigenous microbial communities probably through an antimicrobial effect. Although the available results are still limited, fragmentary and not conclusive, one may assume that plant domestication can lead to changes in root exudation profiles and thereby impact on the rhizosphere microbiome composition and function (Fig. 2).

Effect of plant domestication on belowground interactions

Undermined mycorrhizal symbiosis

The effect of domestication and plant breeding on belowground interactions with soil microorganisms was addressed by pioneering studies with wheat evaluating the ability of ancestors, landraces and modern genotypes to sustain mycorrhizal symbiosis (Kapulnik and Kushnir 1991). The mycorrhizal dependence (MD), i.e. the degree of dependence on mycorrhizal symbiosis for maximum plant growth and yield, was determined for wild and cultivated wheat genotypes. The results showed that a diploid wheat ancestor, *Triticum tauschii,* displayed a higher MD compared to tetraploid or modern hexaploid wheat genotypes (Kapulnik and Kushnir 1991). Hetrick *et al*. (1992) further showed that

ancestors and primitive hexaploid wheat landraces benefitted more from mycorrhizal symbiosis than modern cultivars. Subsequently, Hettrick *et al*. (1993) determined that the ancestral genotype *Triticum tauschii* var. *strangulata,* the donor of the D genotype in hexaploid modern wheat, showed a higher MD as reported in previous studies, whilst AB genome ancestors did not show mycorrhizal dependence. In these studies, the highly fertile conditions used during the plant breeding process were proposed as a possible explanation for the reduced mycorrhizal dependence of modern genotypes. To support this observation, Hetrick *et al*. (1995) further showed that wheat varieties released before 1975 displayed a higher mycorrhizal responsiveness (MR), defined as the effect of the mycorrhizal symbiosis on plant growth as compared to plants without mycorrhiza, while those released after this date were less responsive. Accordingly, Zhu *et al*. (2001) also found a reduction in MR in Australian modern wheat cultivars as compared to old cultivars. However, these findings were recently contrasted in a meta-analysis of mycorrhizal responsiveness in wild and annual crop plants. Lehmann *et al*. (2012) found that newer genotypes were more mycorrhiza-responsive compared to the ancestral genotypes although less intensively colonized. A possible explanation for this observation is that ancient genotypes, and to a larger extent wild relatives, may have developed adaptations to low nutrient environments and are less dependent on mycorrhizal infection than newer genotypes (Koide *et al*., 1988). However, a decrease in the ability to sustain AMF symbiosis in modern cultivars has been also found for other crops. For instance, it was shown that domesticated breadfruit cultivars (*Artocarpus altilis*) were less able to support AMF as compared with wild ancestors as revealed by significant reductions of vesicular and arbuscular colonization (Xing *et al*., 2012). In maize, the response of four landraces and one hybrid to AMF in two different phosphorus (P) regimes was evaluated; two local landraces were significantly more colonized by AMF and acquired more phosphorus in shoots under low and medium P regimes as compared to the modern maize

hybrid. Interestingly, one of the landraces presented an outstanding mycorrhizal colonization and presented the highest percentage increase in root volume under both P regimes (Sangabriel-Conde *et al*., 2014). The diversity of AMF in the roots of the four landraces and the hybrid was assessed through nested PCR of AMF rDNA and it was shown that the landrace with higher mycorrhizal colonization and P acquisition efficiency also presented the highest number of Glomeromycota OTUs (Sangabriel-Conde *et al*., 2015). The authors proposed that the adoption of native landraces of maize may preserve mycorrhizal symbiosis in these agricultural landscapes.

Domestication effect on rhizobia and other microbes

The effect of plant domestication has also been assessed for the symbiosis between legumes and rhizobia. In a study with pea (*Pisum sativum)*, broad bean (*Vicia faba)* and several wild legumes from the genera *Vicia* and *Lathyrus*, grown in a non-agricultural soil, it was shown that the ability to interact with symbionts was limited for pea and broad bean as compared to promiscuous wild legumes that were able to exploit the diverse rhizobial community (Mutch and Young, 2004). Similarly, it was found that *Cicer reticulatum*, ancestor of cultivated chickpea, showed association with a more diverse *Mesorhizobium* population than modern chickpea (Kim *et al*., 2014). In the legume-rhizobia symbiosis, Kiers and Denison (2008) described that plants can sanction less effective symbionts and invest more resources in highly efficient strains. In order to assess whether the ability to sanction non-effective rhizobia strains was also affected in the breeding process, six soybean cultivars representing 60 years of breeding were evaluated in the simultaneous presence of effective and ineffective rhizobia strains. Kiers *et al.* (2007) showed that newer cultivars had less seed yields as compared to older cultivars and also that the yield difference ratio, i.e. the ability of cultivars to reach their full

symbiotic potential in the presence of mixed rhizobial strains, was higher for older cultivars as compared to newer cultivars.

For the effects of plant domestication on other rhizosphere microbes only few examples exist to date. Germida and Siciliano (2001) revealed that the rhizosphere bacterial community of ancient landraces was more diverse than that of two modern cultivars. In the rhizosphere of the ancestral landrace, Pseudomonads were the dominant genus and also higher numbers of *Aureobacter* were found as compared to modern cultivars (Germida and Siciliano 2001). Also in maize, the influence of its progenitor *Zea mays* subsp. *parviglumis* (Balsas teosinte) and two domesticated maize cultivars on the rhizosphere bacterial and fungal community composition was evaluated (Szoboszlay *et al*., 2015). Shannon's and Simpson's diversity indices for bacterial T-RLFP profiles were higher for teosinte compared with one domesticated cultivar and the same as the other cultivar and the control. Interestingly, the same domesticated cultivar with lower bacterial diversity also showed a lower fungal diversity compared with bulk soil controls (Szoboszlay *et al*., 2015).

Domestication and changes in root architecture

Differences in root architecture between modern cultivars and their wild relatives have been described for a number of crops. For instance, cultivated lettuce (*Lactuca sativa*) produced a shallower root system compared with wild lettuce (*Lactuca serriola*). In cultivated lettuce an inadvertent selection of more laterals roots at the top of the tap root helps plants to respond to surface application of water and fertilizer in crop fields, whereas wild lettuce showed a root system able to access deeper portions of soil (Jackson, 1995). Changes in root architecture have also been described for drought tolerant plant cultivars. For example, a drought tolerant accession of wild barley presented different root length, root dry weight and root volume compared with a modern cultivar, both under control

and drought conditions (Naz *et al*., 2012; 2014). Similarly, the teosinte *Zea mays* subsp. *parviglumis* showed a higher root to shoot dry weight ratio and a higher number of very fine and thick roots than two domesticated maize cultivars. Although the Shannon's and Simpson's diversity indices for the bacterial communities were higher for teosinte compared with one the domesticated cultivar, the relative contribution of the root architecture for the observed microbiome differences was not investigated (Szoboszlay *et al*., 2015). It has been postulated that changes in root architecture due to breeding process may have an effect on the rhizosphere microbiome (Micallef *et al*., 2009b), however, more detailed studies will be needed to investigate this.

Reinstating beneficial partnerships in modern crop cultivars

Over the past decades, plant breeders have exploited genes from native relatives of modern crop species to improve plant growth and health. For instance, wild relatives have been used as sources of alleles to improve the ability of modern cultivars to withstand biotic and abiotic stresses in wheat (Nevo *et al*., 2010; Budak *et al*., 2013; Placido *et al*., 2013), barley (Schmalenbach *et al*., 2008) and lettuce (Johnson *et al*., 2000; Simko *et al*., 2013). Similarly, entomologists have explored native habitats to identify natural enemies of insect pests. In the area of plant microbiome research, relatively few efforts have been made to study the biodiversity and functions of beneficial microbial communities present in the native habitats of ancestors of modern crop species. In a study comparing the microbiome of sugar beet and its ancestor *Beta vulgaris* spp. *maritime*, plants were grown in agricultural and in native soils (natural habitat). Wild beet plants showed a more diverse bacterial community compared with domesticated sugar beet, as was shown by single strand conformation polymorphism (SSCP) analysis of the 16S-rRNA genes from total community DNA and 16S amplicon pyrosequencing (Zachow *et al*., 2014). A first approach to identify plant loci associated with root colonization and pathogen protection

by beneficial microorganisms involved a study with six inbred tomato lines and the biocontrol bacterium *Bacillus cereus* UW85 (Smith *et al*., 1997). Based on a dose-response model, they found differences for both intrinsic plant resistance to pathogen infection and support of biological control in the tomato lines. In a follow-up study with several recombinant inbred line (RIL) populations derived from an interspecific cross of cultivated tomato and the related wild species *Lycopersicon cheesmanii*, Smith *et al*. (1999) showed that several quantitative trait loci (QTL) were associated with support of growth of the biocontrol agent *B. cereus* UW85 and the disease suppressive effects. However, several conceptual and experimental efforts have yet to be made in order to identify and exploit these traits in a rhizosphere-based breeding program (Bakker *et al*., 2012). The search for plant traits linked to microbial recruitment by wild relatives holds a huge potential to elucidate and exploit beneficial interactions between plants and microbes. This hypothesis relies on the assumption that wild plant relatives have coevolved with the microbial community of native soils, performing an active selection of microbes with beneficial effects on plant growth and health. We hypothesize that wild relatives are able to establish, with higher frequency, beneficial interactions with microbes as compared to domesticated cultivars. In this context, we propose a pipeline for this emerging research area (Fig. 3).

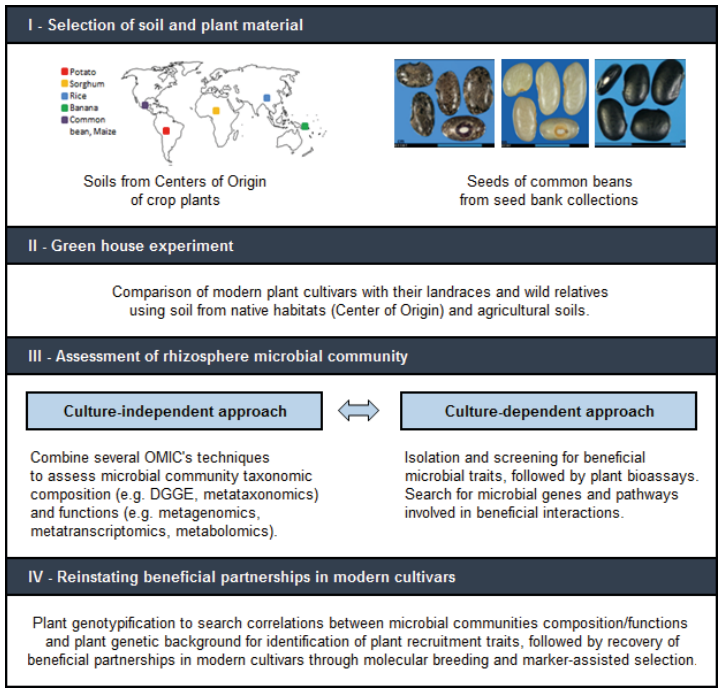


Fig. 3. General workflow to investigate the possibility to reinstating beneficial partnerships in modern cultivars by assessing the rhizosphere microbiome of wild plants in native soil.

First, it is necessary to know the evolutionary history and process of domestication of the host plant to make a proper selection of wild plant materials, as well landraces and modern cultivars. If possible, the modern cultivars should be derived from the selected landraces; however this is not possible for all cultivated species, where the full domestication trajectory is unknown. In parallel, the centre of origin and centres of diversification should be known. Botanic and archaeological records have been used to determine where the wild relatives of many modern crops were originally formed as a species, followed by the domestication process and possible routes of dispersion by humans. Using this information, the collection of native soils in pristine sites located in the centre of origin and its use in the experimental setup will provide the native microbial assemblage in which wild relatives presumably recruit and sustain a more beneficial microbiome compared with less competent landraces and modern cultivars. To evaluate the impact of

the different plant genotypes on the rhizosphere microbiome composition and functional potential, the use of metagenomics and metatranscriptomics together with culture-dependent approaches can be used to identify shifts in taxonomic and functional diversity of the microbiomes of the different plant genotypes. Based on these "omics" data, a screening with culture-dependent approaches can be performed by targeted isolation of those microbial genera that are specifically or more predominantly recruited by wild plant relatives. Evaluation of antagonistic activities against soil-borne pathogens, nutrient solubilisation or improved drought tolerance of the plant species after introduction will help to pinpoint those beneficial activities that plants look for in microbial partners. Finally, once the recruitment of particular taxa is confirmed and the utility of this association is determined, a plant genotyping strategy, quantitative trait loci (QTL) mapping and genome wide association studies (GWAS) with wild relatives, landraces, modern cultivars and preferably crosses between these plant genotypes must be performed in order to identify specific regions in the genome where the recruitment traits are located. Consequently, molecular breeding and marker-assisted selection can be applied to improve beneficial plant-microbe interactions in crop systems. Therefore, this approach of 'going back to the roots', i.e. assessing and accessing the microbiome of indigenous plants and their native habitats, represents a yet untapped avenue to further exploit microbes and plant traits in modern agriculture.

Chapter 3

Linking rhizosphere microbiome composition of wild and domesticated *Phaseolus vulgaris* to genotypic and root phenotypic traits

Juan E Pérez-Jaramillo, Victor J Carrión, Mirte Bosse, Luiz FV Ferrão, Mattias de Hollander, Antonio AF Garcia, Camilo A Ramírez, Rodrigo Mendes and Jos M Raaijmakers

The ISME Journal 11(10):2244-2257 (2017)

https:// doi: 10.1038/ismej.2017.85

Abstract

Plant domestication was a pivotal accomplishment in human history, but also led to a reduction in genetic diversity of crop species compared to their wild ancestors. How this reduced genetic diversity affected plant-microbe interactions belowground is largely unknown. Here, we investigated the genetic relatedness, root phenotypic traits and rhizobacterial community composition of modern and wild accessions of common bean (*Phaseolus vulgaris*) grown in agricultural soil from the highlands of Colombia, one of the centres of common bean diversification. DArT-based genotyping and phenotyping of local common bean accessions showed significant genetic and root architectural differences between wild and modern accessions, with a higher specific root length for the wild accessions. Canonical Correspondence Analysis indicated that the divergence in rhizobacterial community composition between wild and modern bean accessions is associated with differences in specific root length. Along the bean genotypic trajectory, going from wild to modern, we observed a gradual decrease in relative abundance of Bacteroidetes, mainly *Chitinophagaceae* and *Cytophagaceae*, and an increase in relative abundance of Actinobacteria and Proteobacteria, in particular *Nocardioidaceae* and *Rhizobiaceae*, respectively. Collectively, these results establish a link between common bean domestication, specific root morphological traits and rhizobacterial community assembly.

Keywords: common bean, root phenotyping, rhizosphere microbiome, plant domestication, wild relatives.

Introduction

The rhizosphere microbiome has a profound impact on plant health and growth by providing key functions involved in nutrient acquisition, abiotic stress tolerance and protection against pathogen infection (Mendes *et al*., 2011; 2013; Bulgarelli *et al*., 2013). Edaphic factors and plant genotype shape, to a certain extent, the composition and metabolic activities of the bacterial communities in the rhizosphere (Berg and Smalla, 2009; Bulgarelli *et al*., 2012; Lundberg *et al*., 2012; Philippot *et al*., 2013). The effects of the plant genotype on rhizosphere microbiome composition has been proposed to be, at least in part, mediated by quantitative and qualitative differences in root exudate composition (Lakshmanan *et al*., 2012; Badri *et al*., 2013; Carvalhais *et al*., 2013; Lebeis *et al*., 2015). Hence, the composition of a particular rhizosphere microbial assemblage is dependent on the plant species (Turner *et al*., 2013; Ofek *et al*., 2014) and even on the cultivar of a given plant species (Peiffer *et al*., 2013).

Plant domestication was essential to human history but also resulted in a significant reduction in genetic diversity of crop species as compared to their wild ancestors (Doebley *et al*., 2006). Whether this reduction in genetic diversity affected specific root morphological traits and microbial diversity and activity in the rhizosphere is still largely unknown. To date, a limited number of studies have indicated that rhizosphere microbiome assembly may have been affected in modern cultivars of plants as compared to their wild ancestors (Bulgarelli *et al*., 2015; Pérez-Jaramillo *et al*., 2016; Leff *et al*., 2016). In this context, wild relatives and also landraces have been proposed to provide valuable new insight into plant traits and genes associated with microbiome assembly, allowing an integral role of microbiome assembly in future plant breeding programs. For most economically important food crops, however, little knowledge is available on the impact of plant domestication on root traits and rhizosphere microbiome assembly. Here, we determined the genetic relatedness and root morphological traits of wild and modern

accessions of common bean (*Phaseolus vulgaris*) and analyzed their rhizosphere microbiome composition. Common bean is the most important legume crop for low-income farmers in Latin America and Africa (Broughton *et al*., 2003; Akibode and Maredia, 2011). Wild common bean originated in central Mexico and spread throughout Central and South America (Gepts, 1998; Bitocchi *et al*., 2012; Desiderio *et al*., 2013). This wide distribution led to geographical isolation of wild common bean and resulted in well characterized genetic pools (Gepts and Bliss, 1985). A vast collection of available accessions, ranging from wild relatives to highly productive modern varieties, makes common bean a good model system to investigate the impact of domestication on root phenotypic traits and on rhizobacterial community composition of an economically important food crop. Furthermore, common bean and other leguminous plant species provide excellent experimental systems to study the intertwined relationships between nodulation and rhizosphere microbiome assembly (Zgadzaj *et al*., 2016).

In this study, we adopted the approach of 'going back to the roots' (Pérez-Jaramillo *et al*., 2016) and selected eight Colombian accessions of common bean, including wild relatives, landraces and modern cultivars and characterized their genetic relatedness by Diversity Array Technology (DArT) (Jaccoud *et al*., 2001). Subsequently, the selected bean accessions were grown in agricultural soil collected from the highlands of Antioquia, Colombia. Colombian mountains are considered an important centre of common bean diversification where wild and landraces of common bean from the two main genetic pools (Mesoamerican and Andean) can still be found in their natural habitats (Gepts and Bliss, 1986). The selected common bean accessions were subjected to phenotypic analyses of different root traits as well as rhizobacterial community analyses by 16S rRNA amplicon sequencing.

Materials and Methods

Selection of common bean accessions and Colombian soil

Two wild, three landraces and three improved varieties (modern cultivars) of common bean (*Phaseolus vulgaris*) were selected based on the following characteristics: they belong to the Colombian Mesoamerican genetic pool; landraces and modern accessions are the same race, they exhibit the Mesoamerican phaseolin protein type; they originate from the same altitudinal range; and they have the same growth type (i.e. climbing instead of bushy). The latter characteristic is the case for all selected accessions, except for accession G5773, which is a bushy commercial variety widely distributed and commonly used in Latin America and Africa. The seeds were kindly provided by the Genetic Resources Program at the International Centre for Tropical Agriculture - CIAT – in Palmira, Colombia. The plant passport is given in Table S1. The soil used in this study was collected from an agricultural field in the rural area of the municipality of El Carmen de Viboral (Antioquia - Colombia, 6°4'55'' N, 75º20'3''W). Common bean has been cropped in this region for decades and soil conditions are optimal for the growth of several common bean varieties. The soil was collected at 3 random sites in the field from a depth up to 30-cm, air dried, passed through a 2-mm mesh sieve to remove (plant) debris and stored for further use. Physicochemical analyses were performed in the Soil Science Laboratory from the National University of Colombia in Medellín, using standard procedures (Table S2).

Genotyping of common bean accessions

The bean seeds were surface-sterilized and germinated on filter paper wetted with sterile tap water. After 2-5 days, germinated seeds were transferred to 500-ml pots filled with agricultural soil. For each bean accession, two seedlings were transplanted to a pot (1 pot per accession), arranged randomly in a growth chamber (25ºC, 16h daylight) and watered

every day. After 10 days, the youngest leaf of each plant was collected and DNA was isolated with the PowerPlant® Pro DNA Isolation Kit (Mo Bio Laboratories, Carlsbad, CA, USA). The manufacturer's instructions were followed and the yield and quality of the DNA was assessed via agarose gel electrophoresis and a Qubit 2.0 fluorimeter (Invitrogen, Life technologies). Genome profiling was performed using the complexity genome reduction method developed by Diversity Arrays Technology Pty Ltd (DArT P/L, Bruce, Australia) (Jaccoud *et al*., 2001). A proprietary analytical pipeline developed by DArT P/L was used to produce single nucleotide polymorphisms (SNP) tables; in total 10,732 SNPs were obtained. The SNP calling was performed using a custom R script (R Core Team, 2015) and after filtering, a total of 7,527 SNPs were retained for further analysis.

Plant genetic diversity

An identity-by-state (IBS) distance matrix was constructed in PLINK (v. 1.9) (Chang *et al*., 2015) and a neighbor-joining phylogenetic tree was created using the Phylip software package (v.3.695). For the quantitative assessment of the number of groups in the panel, a Bayesian clustering analysis was performed using the model based approach implemented in the STRUCTURE software (Pritchard *et al*., 2000). This approach uses multi-locus genotypic data to assign individuals to clusters or groups (K) without prior knowledge of their population affinities and assumes loci in Hardy-Weinberg equilibrium. The software was ran considering K-values ranging from 1 to 6 (hypothetical number of groups) with an admixture model with correlated allele frequencies. Each run was implemented with 20,000 burnin iterations followed by 200,000 MCMC (Markov Chain Monte Carlo) iterations for accurate parameter estimates. Five independent runs for each K were performed. The number of genetic groups was estimated using the STRUCTURE HARVESTER software (Earl, 2012), by the Evanno criterion (Evanno *et*

al., 2005). A multidimensional scaling analysis was also performed using PLINK. The inbreeding coefficient and occurrence of homozygous segments were computed using the commands '*–het*' and '*—homozyg*' in PLINK. The number of homozygous regions as well as their genomic locations was determined for each bean accession. Similarity of bean accession G51283K1 to the other accessions was determined by computing pairwise IBS. The genome was divided into 109 blocks and within each block pairwise IBS was calculated for all bean accessions; zero is completely different and two is completely identical. The accession G51283K1 was compared with the whole genomes of G22304 and landrace G23998 as wild accessions, with modern accessions G5773 and G51695 as modern accessions and with landrace G50632I1. All the genetic diversity and homozygosity analyses were performed in PLINK (v1.9) and visualized in R.

Root morphology

Seeds were germinated as described above and transferred to 3L pots filled with the agricultural soil described above. Three plants per genotype were used. The plants were grown under ambient environmental conditions, with an average temperature of 25ºC and 12h of daylight. When the V4 stage (3rd trifoliate leaf) was reached, the plants were carefully harvested and the root system was gently washed with tap water until no more soil particles were attached to the roots. Subsequently the entire root system was dyed with methylene blue, laid out on a Scanjet G4050 Scanner (Hewlett-Packard, USA) and scanned with a resolution of 600dpi. The images were then analyzed with the software WinRHIZO (Regent Instruments Inc., Canada), and several root measurements were recorded (Table S3). After scanning, roots were dried and root dry weight (rdw) measured. Subsequently, we computed the Specific Root Length (SRL) using the equation rl/rdw, and the Root Tissue Density (D), using the equation rdw/rv (Martin-Robles *et al*., 2015). These parameters were calculated, normality and homogeneity of

variances were checked using Shapiro-Wilk test and Levene's test, respectively, and one way ANOVA and post hoc tests were used to assess differences in root morphology between the bean accessions.

Rhizospheric soil collection and DNA isolation

The same procedure for seed germination described above was followed. Seedlings were transferred to 3L pots containing the agricultural soil. For each accession, four replicates were used with one plant per replicate pot. The plants were arranged randomly in a greenhouse with ambient environmental conditions with an average temperature of 25ºC and 12h of daylight. Four pots with the same amount of soil but without plants were used as controls and served as bulk soil samples. Plants were harvested at flowering to synchronize microbiome analyses for all accessions at the same phenological growth stage. Rhizospheric soil was collected according to the method of Lundberg *et al*. (2012). Briefly, the entire root system was sampled from the pots, soil loosely attached to the roots was removed and subsequently the entire root system was divided in three parts and each was transferred to a 15mL tube containing 5mL of LifeGuard Soil Preservation Solution (Mo Bio Laboratories, Carlsbad, CA, USA). The tubes were vigorously shaken, the roots were removed and at least 1g (wet weight) of rhizospheric soil was recovered per sample for DNA isolation. For the bulk soils, approximately 1g of soil was harvested from each control pot and also submerged in 5mL of LifeGuard solution. Root dry weight, number of days to reach flowering and the total numbers of nodules per root system were scored. To obtain rhizospheric DNA, a RNA PowerSoil® DNA Elution Accessory Kit (Mo Bio Laboratories, Carlsbad, CA, USA) was used according to manufacturer's instructions after a previous step for RNA extraction and elution with a RNA PowerSoil® Total RNA Isolation Kit (Mo Bio Laboratories, Carlsbad, CA, USA). Each obtained DNA sample was then cleaned with the PowerClean® DNA Clean-Up Kit (Mo Bio

Laboratories, Carlsbad, CA, USA). Agarose gel electrophoresis and a ND1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) were used to check DNA yield and quality. DNA samples were stored at -80°C until further use.

16S rRNA amplicon sequencing

The DNA extracted from the rhizosphere was used for amplification and sequencing of the 16S rRNA, targeting the variable V3-V4 regions (16S Amplicon PCR Forward Primer=5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGC WGCAG;16S Amplicon PCR Reverse Primer = 5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCT AATC) resulting in amplicons of approximately ~550 bp. Dual indices and Illumina sequencing adapters using the Nextera XT Index Kit were attached to the V3-V4 amplicons. Subsequently, library quantification, normalization and pooling were performed and MiSeq v3 reagent kits were used to finally load the samples for MiSeq sequencing. For more info please refer to the guidelines of Illumina MiSeq System (Illumina, 2013).

The RDP extension to PANDASeq (Masella *et al*., 2012), named Assembler (Cole *et al*., 2014), was used to merge paired-end reads with a minimum overlap of 10 bp and at least a Phred score of 25. Primer sequences were removed from the per sample FASTQ files using Flexbar version 2.5 (Dodt *et al*., 2012). Sequences were converted to FASTA format and concatenated into a single file. All reads were clustered into operational taxonomic units (OTUs) using the UPARSE strategy by de-replication, sorting by abundance with at least two sequences and clustering using the UCLUST smallmem algorithm (Edgar, 2010). These steps were performed with VSEARCH version 1.0.10 (Rognes *et al*., 2015), which is an open-source and 64-bit multi-threaded compatible alternative to USEARCH. Next, chimeric sequences were detected using the UCHIME

algorithm implemented in VSEARCH (Edgar *et al*., 2011). All reads before the dereplication step were mapped to OTUs using the usearch_global method implemented in VSEARCH to create an OTU table and converted to BIOM-Format 1.3.1 (McDonald *et al*., 2012). Finally, taxonomic information for each OTU was added to the BIOM file by using the RDP Classifier version 2.10 (Cole *et al*., 2014). All steps were implemented in a Snakemake workflow (Köster and Rahmann, 2012). The OTU table was filtered using QIIME (1.9.1) custom scripts (Kuczynski *et al*., 2012). The Bacteria domain was extracted using the command *split_otu_table_by_taxonomy.py* and singletons, doubletons and chloroplast sequences were discarded with the command *filter_otus_from_otu_table.py,* obtaining a filtered OTU table for further analysis.

Rhizobacterial diversity and link with genotypic and root phenotypic traits

The alpha diversity was calculated using QIIME customs scripts. The command *alpha_rarefaction.py* was used to rarefy the OTU table to counts up to 50,000 reads. This was the lowest sequencing depth obtained from a sample and therefore used as a threshold for rarefaction and alpha diversity calculations (Gotelli and Colwell, 2001). The *alpha_diversity.py* command was applied to rarefied data and observed OTUs, Shannon, Chao1 and Faith's Phylogenetic Diversity metrics were obtained. One-way ANOVA and Tukey HSD were performed in R. For the Beta-diversity calculations, the entire filtered OTU table was used and normalized using the function *cumNorm* from the R package metagenomeSeq (v.1.12) (Paulson *et al*., 2016). We used a cumulative-sum scaling (CSS) method, which calculates the scaling factors equal to the sum of counts up to a particular quantile to normalize the read counts in order to avoid the biases generated with current sequencing technologies due to uneven sequencing depth (Paulson *et al*., 2013). Bray-Curtis dissimilarity matrix was calculated and used it to build Principal Coordinate Analyses and Constrained Principal Coordinate Analysis (CAP) constrained by

phylogenetic group, i.e., ancestral (A1 and A2) and modern (M1 to M5), using the function *capscale* retrieved from Vegan package (Oksanen *et al*., 2016) (v.2.3-2) and implemented in the Phyloseq package (McMurdie and Holmes, 2013) (v.1.10), both in R. The nonparametric *adonis* test was used to assess the percentage of variation explained by the Phylogenetic grouping along with its statistical significance. Permutational multivariate analyses of variance were performed to evaluate the significance of the constrained principal coordinate analyses, both retrieved from Phyloseq and Vegan packages. A Regularized Canonical Correlation Analysis was also performed in order to graphically depict whether the genetic make-up of the bean accessions correlates with their rhizobacterial community structures, using the R package CCA (González and Déjean, 2012). The function rcc was used, which is an extension of the Canonical Correlation Analysis to seek correlations between two data matrices. Subsequently, the function plt.cc was used to generate the plots. Canonical Correspondence Analysis (CCA, Canoco 5.0) was also conducted with a complete set of CSS normalized counts of the 16S rRNA data and the root morphological traits specific root length (SRL), root density, root dry weight and number of nodules. The adjusted explained variation was determined, unrestricted permutations were calculated to determine the significant contribution of each variable and Bonferroni corrections were applied to adjust the *P*-values. Constrained ordinations were built using rhizobacterial phyla or families together with plant genotypic and root morphological traits as explanatory variables.

Species abundance distribution and differential abundance analysis

Species abundance distribution models were used to determine whether neutral or niche-based mechanisms were governing the bacterial assembly. We used the command Radfit from the R package Vegan to evaluate broken stick, pre-emption, log-normal, Zipf and Zipf–Mandelbrot rank abundance models and a zero-sum multinomial (ZSM) model

using the TeTame2 software (Jabot *et al*., 2008). The comparison of the models fit was done based on the Akaike Information Criterion (AIC) using the equation AIC = -2 log-likehood + 2 x npar (Mendes *et al*., 2014). AIC values were compared, being the lowest selected as the best fit to the data (Dumbrell *et al*., 2010). To compare the differences in taxonomic composition and to assess whether some bacterial taxa were differentially abundant, we conducted a three step analysis in which we assessed separately the read counts based on Phylum, Family and OTU level. For Phylum and Family level, custom R commands were used in order to aggregate all the reads according to the level chosen. For the OTU level analysis, the function *calculateEffectiveSamples* from the metagenomeSeq R package was applied to the filtered OTU table and features with less than the average number of effective samples in all features were removed. For the analysis at Phylum, Family and OTU level, we used normalized tables applying the CSS normalization as described above. Then, a Zero-Inflated Gaussian (ZIG) Distribution Mixture Model was applied using the *fitZig* function from metagenomeSeq. With the coefficients from the model, we applied moderated t-tests between accessions using the makeContrasts and eBayes commands retrieved from the R package Limma (v.3.22.7) (Ritchie *et al*., 2015). Obtained *P*-values were adjusted using the Benjamini-Hochberg correction method. Differences in the abundance of taxa between accessions were considered significant when adjusted *P*-values were lower than 0.1 at Phylum and Family level, and 0.05 at OTU level. Volcano plots were built to graphically represent the results of the moderated t-tests using the R package ggplot2 (v.2.0.0) (Wickham *et al*., 2009). To graphically represent the results obtained at Phylum and Family level, a script developed by Bulgarelli *et al*. (2015) was adapted, in which relative abundance of read counts per mil was used, as well as box plot representations using the R package ggplot2. Taxa above 5‰ relative abundance were plotted for Phylum and Family level analysis. Treemap (v.3.7.3) was used to visualize the significantly abundant OTU's, the annotated

taxonomy, the adjusted *P*-value and per mil relative abundance in bubble graphs, in which the size of the bubbles indicates de relative abundance per mil of the raw read counts.

Data access and bioinformatic analyses

The sequence data are deposited at the European Nucleotide Archive (ENA) under accession number PRJEB19467. Data, scripts and codes used for statistical and bioinformatic analyses are available at: https://doi.org/10.5281/zenodo.580027 and https://doi.org/10.5281/zenodo.556538, respectively.

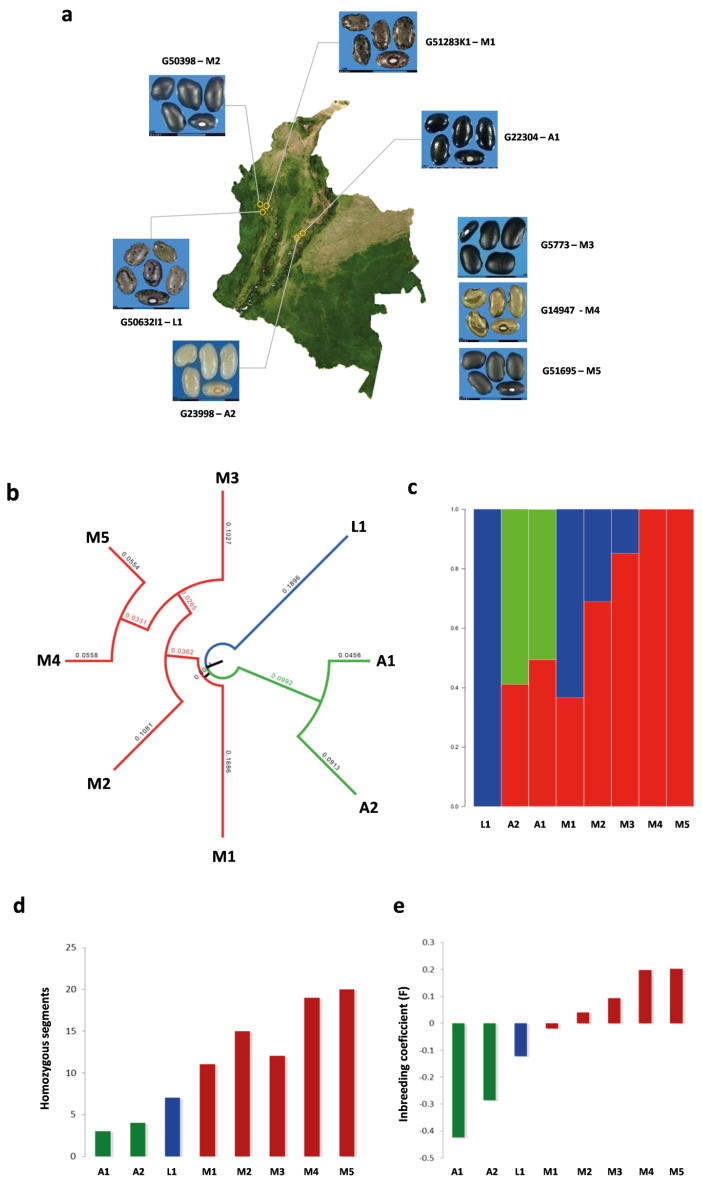


Fig. 1. Origin and genetic structure of the common bean accessions. (a) Map of Colombia depicting the geographic origin, accession number and classification of the bean accessions based on DArT genotyping performed in this study; **(b)** Neighbor joining and phylogenetic relatedness; **(c)** STRUCTURE analysis (k=3); **(d)** Number of homozygous segments, and **(e)** Inbreeding coefficients (*F* values). Green color is assigned to ancestral accessions A1 and A2, blue to the landrace accession L1 and red to modern accessions M1 to M5.

Results and discussion

Genetic relatedness of common bean accessions

DArT analysis resulted in 7,527 SNPs as genetic markers for the eight selected local common bean accessions. Phylogenetic and Bayesian clustering approaches as well as multidimensional scaling allowed us to decipher the divergence among the selected bean accessions (Figs. 1b and c; Fig. S1). Bean accessions G22304 and G23998, originally selected for this study as wild and landrace, respectively, showed strong genetic concordance and were classified as wild or 'Ancestral' accessions A1 and A2, respectively (Fig. 1 and Table S1). For the accession G51283K1, selected originally as a wild, a genome-wide comparison showed that it is more similar to the modern bean accessions than to the wild accessions (Fig. S2). We postulate that G51283K1 is probably a weedy accession, i.e. the product of a cross between wild common bean and a modern cultivar (Toro *et al*., 1990). Hence, accession G51283K1 was classified together with G50398, G14947, G51695 and G5773, as 'Modern'. Hereinafter, these five accessions are referred to as M1-M5, respectively (Fig. 1 and Table S1). Accession G50632I1, selected as a landrace, did not show significant similarity with any of the other bean accessions and was named L1 (Figs. 1b and c; Fig. S3). Inference of the genetic diversity of the selected bean accessions further supported our classification: accessions M1-M5 have a higher number of homozygous segments and higher inbreeding coefficients than landrace L1 and wild accessions A1 and A2 (Figs. 1d and e). Accessions A1 and A2 originate from the same geographic area where several wild relatives of common bean have been collected (Blair *et al*., 2012) (Fig. 1a). The proximity between collection sites might partly explain the genotypic similarities between these two wild accessions.

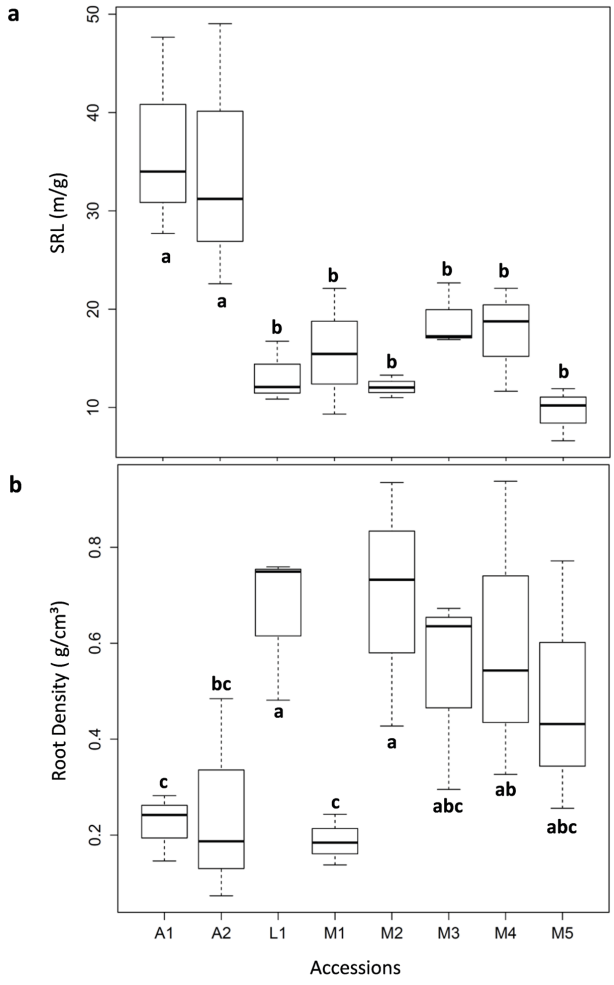


Fig. 2. Root morphology parameters of common bean accessions. (a) Specific Root Length is the product of root length divided by the root dry weight, and **(b)** Root density is the product of root dry weight divided by root volume. Root length and root volume were determined by WinRHIZO. The colors of the lines below the labels of the accessions represent the groups based on their genetic similarity (DArT analysis). Statiscally significant differences between group means for SRL and for Root Density were determined by one-way ANOVA ($P < 0,05$). Three replicates per accession were used. Different letters indicate statistically significant differences (Fisher LSD test).

Root phenotypic traits of wild and modern bean accessions

Wild accessions A1 and A2 had a similar specific root length (SRL: ratio of root length and dry weight) and a similar root density (ratio of dry weight and volume), different

from the other bean accessions (Fig. 2). These results confirm and extend earlier results found for wild common bean as compared to cultivars (Martin-Robles *et al*., 2015). Taken together, a high SRL and small diameter point to thinner roots and may provide a higher efficiency of water search and uptake, traits that are important for wild beans to prosper and survive in the dry native habitats (Toro, 1990; Comas *et al*., 2013). When harvested at flowering stage, significant differences were observed between the bean accessions in the number of days to reach flowering, the root dry weight and the number of nodules (ANOVA, $P<0.005$; Fig. S4). Consistent with previous findings (Toro *et al*., 1990), bean plants of the ancestral group require more days to reach flowering than the modern accessions. Only A2 presented a significant higher number of nodules per root system, while no significant differences were found between the other accessions (Fig. S4c).

The common bean rhizobacterial diversity

As root exudation profiles may change due to specific root architectural features (Marschner *et al*., 2002), we hypothesized that the observed contrasting root morphologies and the genetic divergence between accessions may affect the rhizobacterial community composition. Plants were harvested at flowering to synchronize microbiome analyses for all bean accessions at the same phenological stage. Through sequencing of the V3-V4 region of the 16S rRNA, 2.4 million quality reads were recovered, identifying 12,293 operational taxonomic units (OTUs) at 97% sequence similarity (Table S4). For the α-diversity, we observed a significant reduction in the rhizosphere of all bean accessions as compared to the bulk soil (ANOVA, $P < 0.05$) (Figs. S5 and S6). Between accessions, however, we did not find significant differences in the diversity indexes, except for the number of observed OTUs which was higher for M1 than for A1 (Tukey HSD, $P < 0.05$). Bray-Curtis metrics and Constrained Analyses of Principal Coordinates (CAP) further showed that the microhabitat (soil, rhizosphere)

explained 30.2% of the β–diversity, i.e. the total variability in bacterial community structure between groups (PERMANOVA, $P < 0.001$). Accordingly, a significant separation between rhizosphere and bulk soil was observed (PERMANOVA $P < 0.005$) (Fig. 3a). This selective pressure of the rhizosphere on microbiome composition is well known (Lakshmanan *et al*., 2012; Badri *et al*., 2013) and most likely driven by the quantity and quality of root exudates in combination with different growth rates, substrate utilization spectra and competitive abilities of the rhizobacterial genera. The results further showed that 13.5% of the total variability in rhizobacterial community composition was explained by the bean genotype (PERMANOVA, $P<0.001$). The constrained analysis of the principal coordinates by phylogenetic group was significant (PERMANOVA, $P < 0.005$) (Fig. 3b). A Regularized Canonical Correlation Analysis further confirmed that the genetic make-up of the wild bean accessions correlates with their rhizobacterial community composition (Fig. S7). These results are in accordance with previous findings on maize and barley, where the impact of the plant genotype shapes host-dependent rhizosphere bacterial communities (Peiffer *et al*., 2013; Bulgarelli *et al*., 2015).

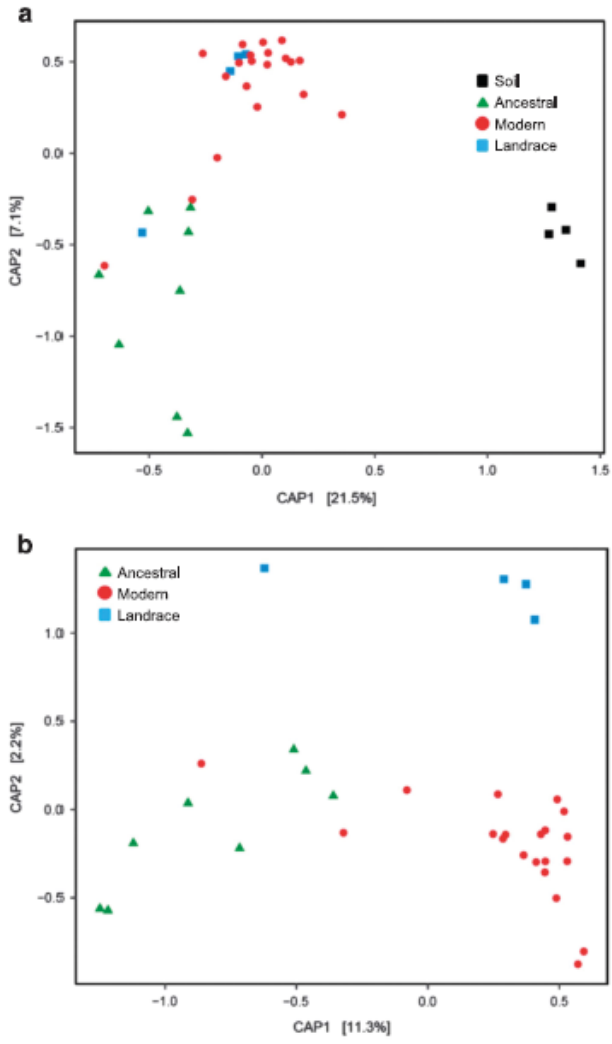


Fig. 3. Rhizosphere bacterial community structure of common bean. Constrained Analysis of Principal Coordinates (CAP) of 16S rRNA diversity in the rhizosphere of the eight common bean accessions used in this study, with **(a)** and without **(b)** 16S rRNA diversity in the bulk soil, respectively. CSS transformed reads were used to calculate Bray-Curtis distances and a constrained analysis was performed by microhabitat **(a)** (30.2% of the overall variance; P<0.005) and bean group **(b)** (13.47% of the overall variance; P<0.005). Statistical significance of the Constrained analysis was assessed by Permanova (P<0.005).

Niche-based processes in rhizobacterial community assembly

The SAD models and the comparison of AIC values showed that the rhizobacterial species abundance in the rhizosphere of all accessions and in the bulk soil are explained by niche-based distributions (Fig. S8 and Table S5). In the case of the rhizosphere

environment, root exudation is a strong modulating factor of rhizobacterial communities, where several taxa can thrive and become highly abundant whilst other community members exhibit low abundance (Jones *et al*., 2009). We also tested a neutral model in order to generate a SAD to be compared with the other niche-based models. With this model, a parameter (*m*) which accounts for the immigration rate into local communities from a regional pool is obtained (Etienne *et al*., 2005). Values closer to 1 indicate no dispersal limitation. The *m* values for bulk soil samples were closer to 1 as compared to those of the rhizosphere samples, suggesting a possible effect of neutral-driven processes in the bulk soil and at the same time a stronger niche-driven process in the rhizosphere of all the bean accessions tested (Table S6).

Linking rhizobacterial community composition with the common bean genotype

To determine which rhizobacterial taxa were affected in a bean genotype-dependent manner, a ZIG model was used to assess the differential abundance. At phylum level, all eight bean accessions presented an enrichment of Proteobacteria and a lower abundance of Acidobacteria as compared to the bulk soil (Fig. S9a). The phyla Bacteroidetes and Verrucomicrobia were significantly more abundant in the rhizosphere of the wild bean accessions, whereas the Actinobacteria were more abundant in the rhizosphere of the modern bean accessions (FDR<0.1; Figs. 4a and b). At family level, a significant increase in the relative abundance of the *Rhizobiaceae* and *Sphingomonadaceae* was observed for all bean accessions as compared to the bulk soil (Fig. S9b). Following the bean genotypic trajectory from A1 thru M5 (based on inbreeding coefficient and homozygosity), we observed a gradual decrease in the relative abundance of *Chitinophagaceae* and *Cytophagaceae*, both of the Bacteroidetes phylum (FDR<0.1; Figs. 4c and d).

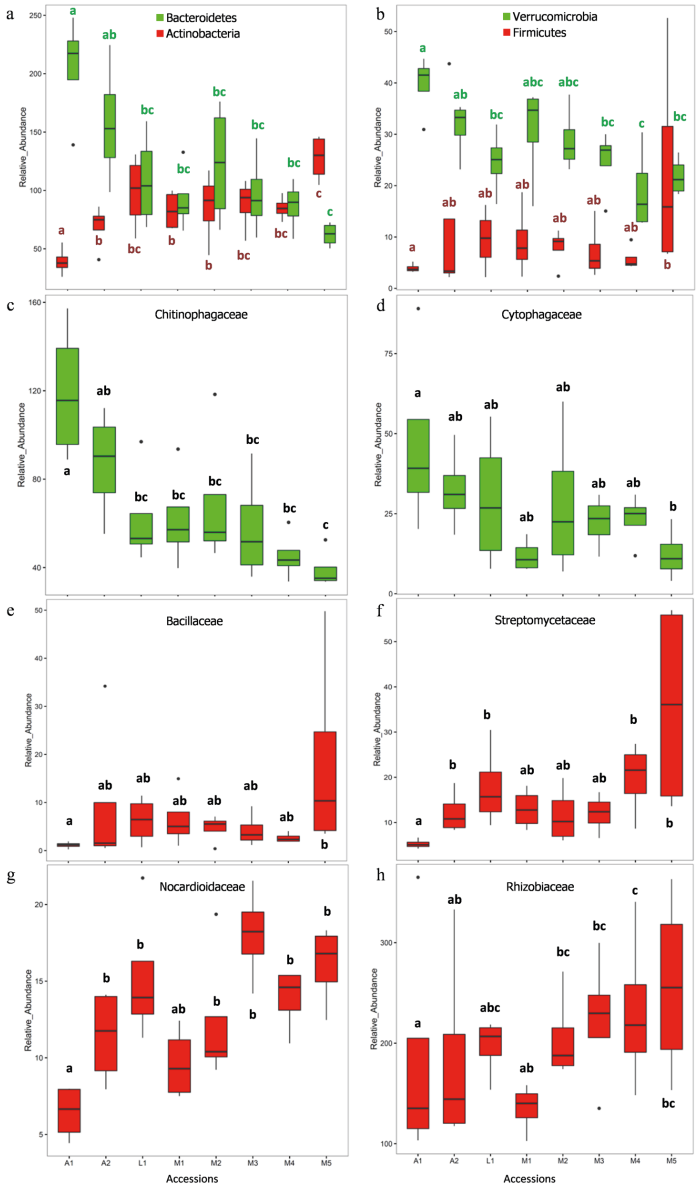


Fig. 4. Relative abundance of bacterial phyla and families in the rhizosphere of the different bean accessions. The relative abundance (‰) of the phyla and families of 4 replicates per accession was used. At phylum level, results are shown for (a) Bacteroidetes and Actinobacteria, and (b) Firmicutes and Verrucomicrobia. At family level, results are shown for representatives of the Bacteroidetes (*Chitinophagaceae* (c) and *Cytophagaceae* (d)). For the Firmicutes and Actinobacteria, results are shown for *Bacillaceae* (e), *Streptomycetaceae* (f) and *Nocardioidaceae* (g). For the phylum Proteobacteria, the relative abundance of the *Rhizobiaceae* (h) is shown. Different letters indicate significant differences between accessions (moderated t test, FDR < 0.1).

Following this same trajectory, gradual increases in relative abundance were observed for the *Nocardiodaceae* (Actinobacteria) and *Rhizobiaceae* (Proteobacteria) (Figs. 4g and h) and to some extent also for the *Streptomycetae* (Fig. 4f). For *Bacillaceae*, however, no specific pattern in relative abundance was observed along the bean genotypic trajectory; this family was significantly more abundant only in the M5 rhizosphere (Fig. 4e). Next, we zoomed in on specific OTUs that were differentially enriched or depleted among the bean accessions by using the filtered OTU table. Based on the inbreeding coefficients and number of homozygous segments (Figs. 1d and e), bean accessions A1 and M5 were the most divergent and therefore compared first to see if this divergence was also reflected in the rhizobacterial community composition. We found 221 OTUs enriched in the A1 rhizosphere and 181 OTUs enriched in the M5 rhizosphere (Fig. 5 and Table S7). A1 was significantly enriched with representatives of the *Chitinophagaceae* family (25 OTUs). The genus *Dyadobacter* from the *Cytophagaceae* family was particularly enriched in the rhizosphere of A1. For the M5 rhizosphere, three out of the ten most abundant OTUs were significantly enriched, belonging to *Rhizobium* (OTU9047), *Streptomyces* (OTU7) and Burkholderiales (OTU8). Also enriched in M5 were 2 highly abundant OTUs of the genus *Arthrobacter* (OTU17 and OTU886) and several OTUs from the family *Nocardioidaceae* (13 OTUs) and the genus *Lysobacter* (8 OTUs). All microbiome comparisons between A1 and the other bean accessions showed similar enrichments (Figs. S10-S14). Collectively, these analyses showed that OTUs from Bacteroidetes and Verrucomicrobia were enriched in the rhizosphere of accession A1, whereas OTUs from Actinobacteria were consistently enriched in the rhizosphere of accession M5. Similarly, when comparing A2 to M4 and to M5, we observed that OTUs from *Chitinophagaceae* family were consistently enriched in the A2 rhizosphere (Figs. S15-S16). Also when we merged the data of the individual bean accessions into a collective data set for each of the two bean genotypic groups (i.e. ancestral and modern), similar overall patterns and differences

in rhizobacterial community composition were observed: Bacteroidetes (8 OTUs) and Verrucomicrobia (5 OTUs) were enriched in the ancestral group, whereas the phylum Actinobacteria (2 OTUs) was enriched in the modern group (Fig. S17).

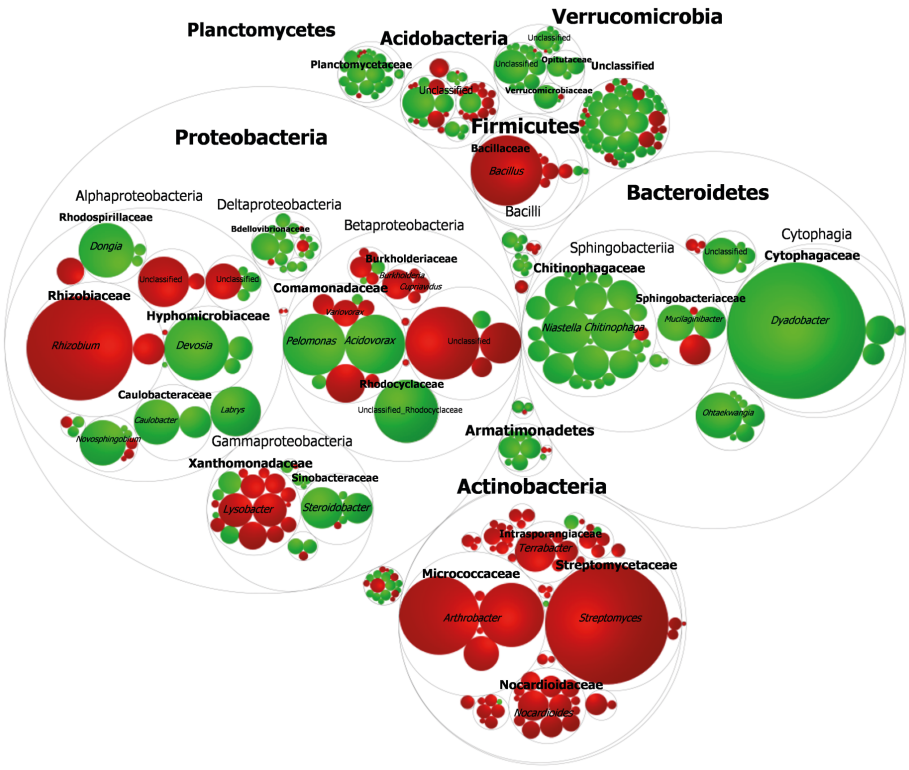


Fig. 5. Differential abundance of bacterial OTUs between the wild bean accession A1 and modern bean accession M5. The comparison was made using a zero-inflated Gaussian distribution mixture model followed by moderated t-test and a Bayesian approach. Data from 4 replicates per accession was used. Only OTUs significantly enriched in one of the two accessions are shown (FDR<0.05). The largest circles represent Phylum level. The inner circles represent Class and Family level. The color of the circles represents the OTUs enriched in the rhizosphere of wild accession A1 (green) or of modern accession M5 (red), with the assigned *Genus* in italics. The size of the circle is the mean read relative abundance of the differentially abundant OTU.

Intriguingly, Bacteroidetes have also been reported at higher relative abundance in the rhizosphere of other wild plant species and wild crop relatives, including *Cardamine hirsuta* (Schlaeppi *et al*., 2013), *Beta vulgaris* subsp. *maritima* (Zachow *et al*., 2014) and

Hordeum vulgare subsp. *spontaneum* (Bulgarelli *et al*., 2015). In human microbiome research, Bacteroidetes have received considerable attention for their association with low carbohydrate diets, lean mice and weight loss in humans (Ley *et al*., 2006; Turnbaugh *et al*., 2006; De Filippo *et al*., 2010; Brown *et al*., 2012). Considering the increased abundance of Bacteroidetes on the thin roots of wild relatives of common bean and the higher relative abundance of Actinobacteria and Proteobacteria on the thicker roots of modern varieties it is tempting to make an analogy of 'lean beans' and 'obese beans'. Whether the association with Bacteroidetes might result in a healthier development for plants as shown for animals, and whether its increased abundance on the roots of the wild relatives is a signature of coevolution remains speculative and opens exciting avenues for further research. More specifically, the putative link between the abundance of this bacterial Phylum in the rhizosphere of wild crop relatives and its ability to degrade complex biopolymers (Thomas *et al*., 2011) will be subject of future experiments. Also in terms of plant health, the *Chitinophagaceae* family, which belongs to the Bacteroidetes, has been proposed for their potential role in protection against soil-borne pathogens (Yin *et al*., 2013; Chapelle *et al*., 2016).

Linking rhizobacterial community composition with root phenotypic traits

Canonical Correspondence Analysis revealed that the overall variation in rhizobacterial community composition is explained for 11.4% (Bonferroni adjusted P=0.002) by the different root phenotypic traits (Fig. 6a), which resembles the percentage of variation (13.5%) in rhizobacterial composition explained by the common bean genotype (Fig. 3b). Amongst the root morphological and phenological traits included in the analysis, the specific root length (SRL) was responsible for most of the explained variation (Bonferroni adjusted P=0.008) followed by the number of nodules (Bonferroni adjusted P=0.016). The percentages of explanation for the variables Root Density and Root Dry Weight were

not statistically significant. Interestingly, the dissimilarities between wild accessions A1 and A2 appears to be largely driven by the number of nodules. When the explanatory variables are used together with the bacterial phyla (Fig. 6b) and bacterial families (Fig. 6c), the results show that the abundance of Bacteroidetes families is explained by SRL, in contrast to the abundance of Actinobacterial families.

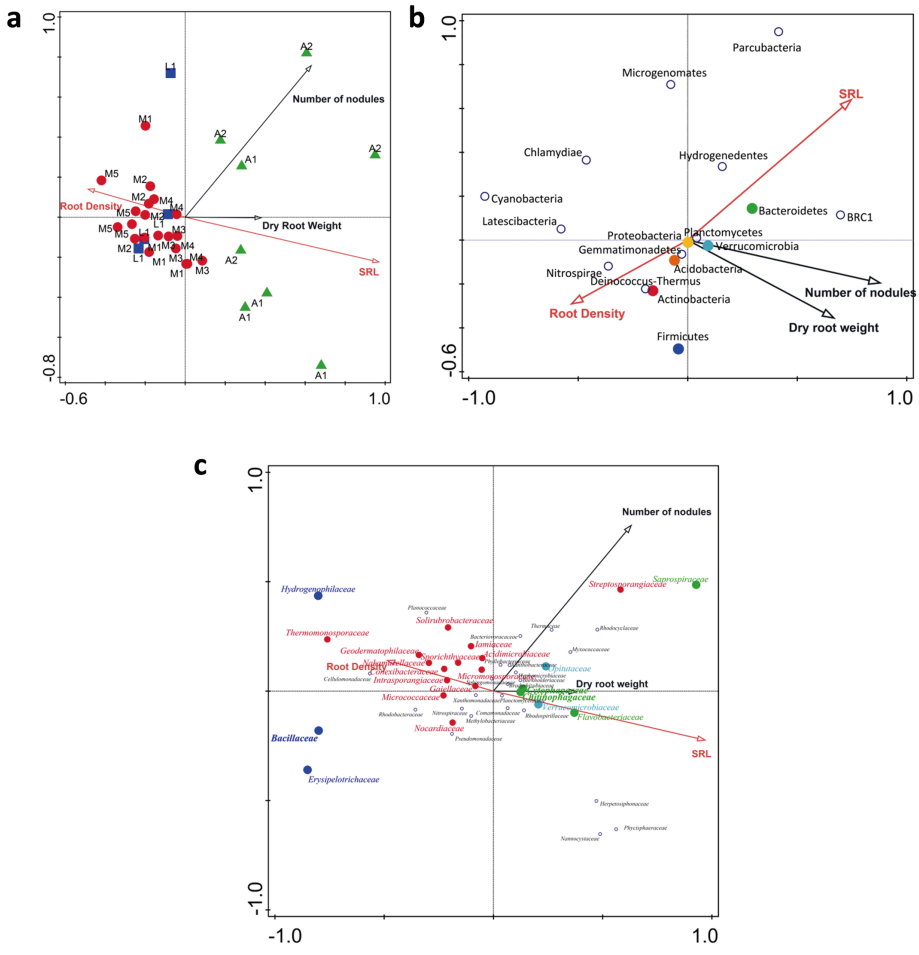


Fig. 6. Constrained Canonical Correspondence Analysis of 16S sequence data and root morphological traits. (a) Root morphological traits as explanatory variables for the divergence between the overall rhizobacterial community composition of the eight different bean accessions. The rhizobacterial community composition is based on CSS normalized counts. Green triangles represent the replicates of the wild bean accessions A1 and A2, blue squares represent the landrace accession L1, and red circles represent the modern bean accessions M1 to M5. The colored arrows represent the root morphological traits: Number of nodules and Dry root weight (black), Specific Root Length (SRL) and Root Density (red). (b) Same as

in panel a depicting the root morphological traits as explanatory variables for the divergence between the different bacterial phyla. Only phyla with a relative abundance higher than 1% are colored: Acidobacteria (orange), Actinobacteria (red), Bacteroidetes (green), Firmicutes (blue), Proteobacteria (yellow) and Verrucomicrobia (light blue). (c) Same as in panel b, depicting the root morphological traits as explanatory variables for the divergence between the bacterial families. Here, only the bacterial families belonging to Actinobacteria (red), Bacteroidetes (green), Verrucomicrobia (light blue) and Firmicutes (blue) were highlighted.

Conclusions

In this study, we found significant associations between the rhizobacterial community composition, the common bean genotype and specific root phenotypic traits. The phyla Bacteroidetes and Verrucomicrobia were consistently more abundant in the rhizosphere of wild common bean accessions, whereas representatives of the phyla Actinobacteria and Proteobacteria were enriched on roots of modern bean accessions. What the impact of the observed shifts in microbiome composition is on growth and health of common bean will be subject of future studies, ultimately providing an answer to the larger question if plant domestication compromised (or not) the beneficial effects of the rhizosphere microbiome. The divergence in rhizobacterial community composition between wild and modern bean accessions suggest a plant genetic basis of rhizosphere microbiome assembly. While these concepts apply also to other important food crops (e.g., cereals), only with legumes it is possible to study how nodulation, the rhizosphere microbiota and the relationships between these two types of plant-microbe interactions are intertwined. In our study, only wild bean accession A2 presented a higher number of nodules per root system, while no significant differences were found between the other bean accessions. This suggests that symbiotic nitrogen fixation per se may not be the major driver of the root microbiome composition as was elegantly shown recently for *Lotus japonicum* (Zgadzaj *et al*., 2016). These results also imply that other or additional host-derived cues shape the bean rhizosphere microbiota. The relatively small sample size used in our study precludes a statistically robust GWAS analysis but did provide a well

differentiated set of traits in wild and modern accessions associated with a number of bacterial taxa. In-depth genetic and phenotypic analyses of a larger population of plant accessions (Kraft *et al*., 2010) will be needed for the identification of genes or molecular markers that ultimately can be used in plant breeding programs for the recruitment of specific plant-beneficial microbial taxa.

Acknowledgements

J.E.P-J was financially supported by the Department of Science, Technology and Innovation of Colombia - COLCIENCIAS through the doctoral grant 568-2012-15517825. J.M.R. and V.J.C. were supported by the Dutch STW-program "Back to the Roots" and R.M. by CNPq 443112/2014-2. We are grateful for the expert help from Dr. O. Toro (R.I.P.), Dr. D. DeBouck and L.G. Santos at the International Centre for Tropical Agriculture (CIAT, Cali, Colombia) in the selection of the bean accessions. We are also grateful to H. A. Pérez and J. A. Osorio for collecting the soil for the greenhouse experiments. We thank G. Bongiorno and N.M. Robles for their valuable help in root phenotyping and J.N. Paulson for his expert advice in using the R package metagenomeSeq.

Supplementary materials

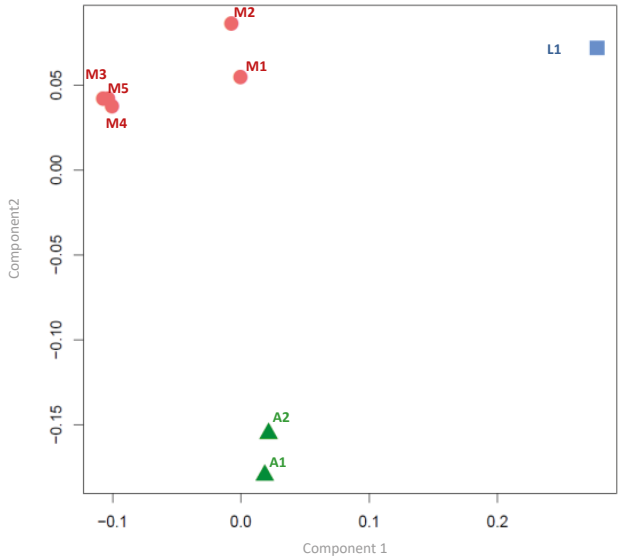


Fig. S1. Multidimensional scaling of common bean accessions. MDS is based on IBS distance matrix analysis. Green color is assigned to the wild bean accessions (A1 and A2); red color is assigned to modern accessions (M1-M5); and blue color is assigned to Landrace L1.

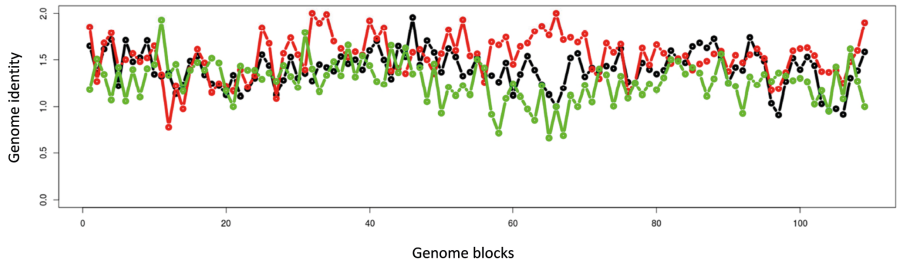


Fig. S2. Genome comparisons between bean accession 51283K1 (M1), landrace L1 and wild accessions A1 and A2. The bean genome was divided into 109 blocks (the x-axis, ordered according to chromosomes and position). Then it was calculated for each block how similar the genotypes between two accessions are: 0 is different, 2 is identical. The red line indicates how similar the accession 51283K1 (M1) is to two modern accessions M4 and M5. The black line indicates how similar the accession M1 is to the wild accessions A1 and A2; the green line represents the similarity of M1 to landrace accession L1. For most of the genome, M1 is more similar to the modern bean accessions.

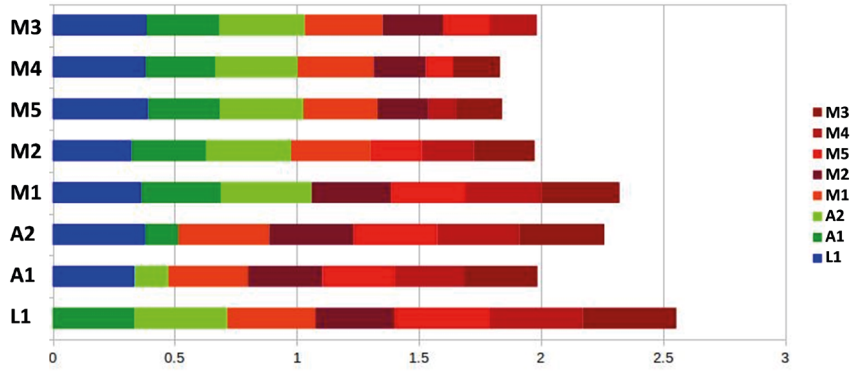


Fig. S3. Divergence between the common bean accessions. Molecular differentiation between accessions, ranging from 0 (identical) to 1 (different for each SNP) in pairwise comparisons. The horizontal axis shows the overall length of the accumulated differences of one accession with all the other accessions, giving a graphical indication of how unique a particular accession is.

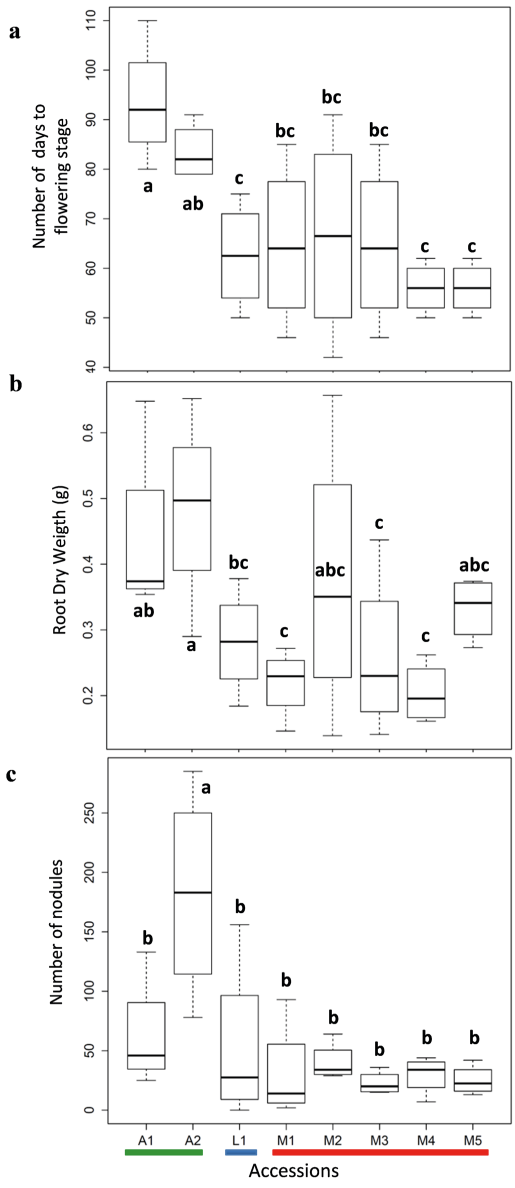


Fig. S4. Phenological traits of the common bean accessions grown in agricultural soil. (a) Number of days to reach flowering, (b) root dry weight, and (c) number of nodules per plant. The colors of the lines below the labels of the accessions represent the groups based on their genetic similarity (DArT analysis). Four replicates per accession were used. Statistically significant differences were found between group means as determined by one-way ANOVA ($P < 0.05$). Different letters denote statistically significant differences between accessions according to Fisher LSD.

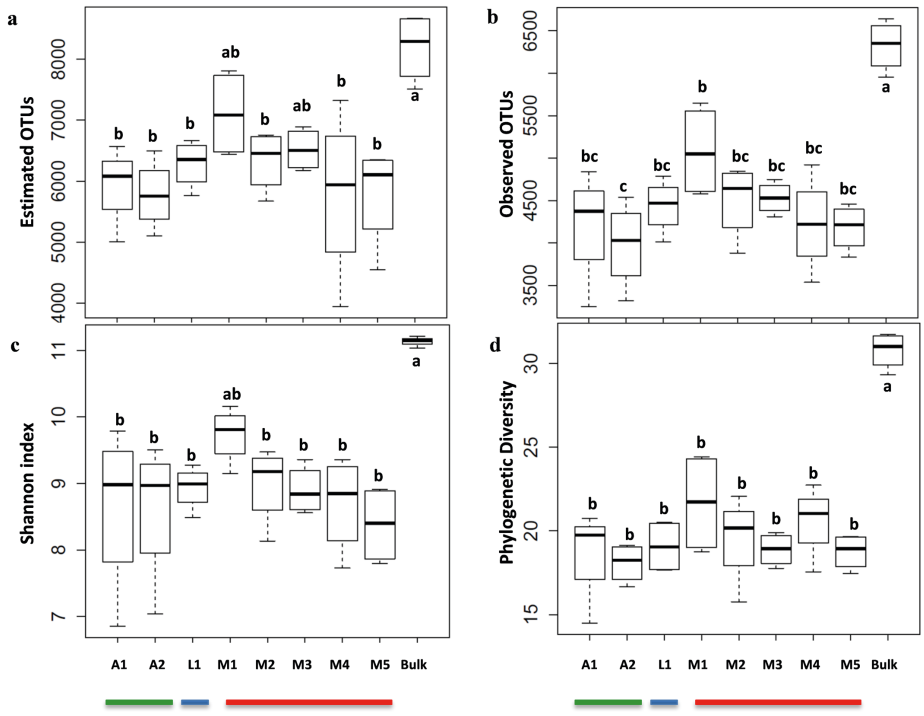


Fig. S5. Comparative analysis of the Alpha diversity of 16S sequence data of the rhizosphere of common bean accessions and bulk soil. (a) Chao1; (b) Observed OTUs; (c) Shannon and (d) Phylogenetic diversity for all bean accessions and the bulk soil rarefied to 50,000 counts per sample. The colors of the lines below the labels of the accessions represent the groups based on their genetic similarity (DArT analysis). There were statistically significant differences between the means as determined by one-way ANOVA ($P<0.05$). Different letters denote statistically significant differences according to Tukey HSD test. The bulk soil is, in general, more diverse than the rhizosphere microbiome of the bean accessions.

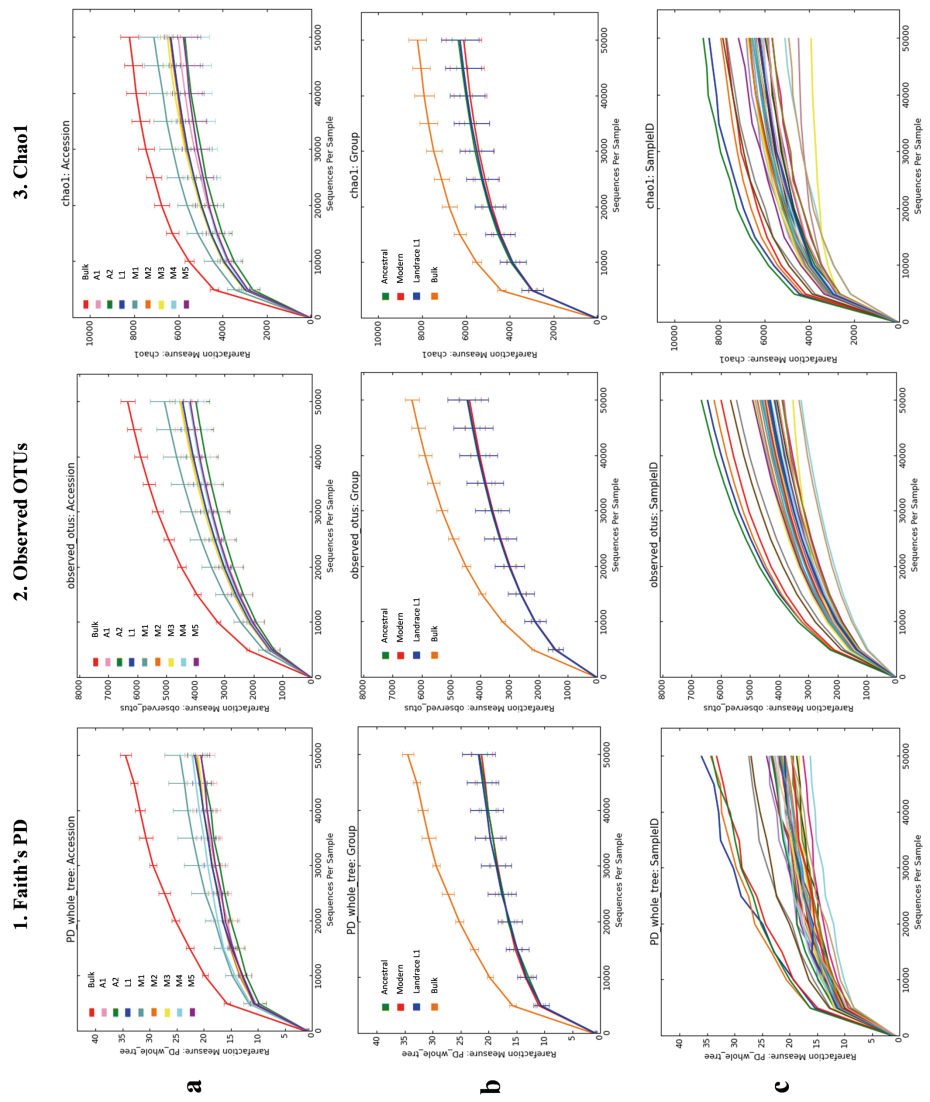


Fig. S6. Rarefaction curves for bulk soil and the bean accessions. In the columns, the rarefaction curves are shown by index as follows: 1. Faith's phylogenetic diversity; 2. Number of observed OTUs; 3. chao1 index for estimated OTUs. In the rows, rarefaction curves are shown by (a) Bean accession; (b) Bean group classification according to DArT analysis, and (c) by sample replicate.

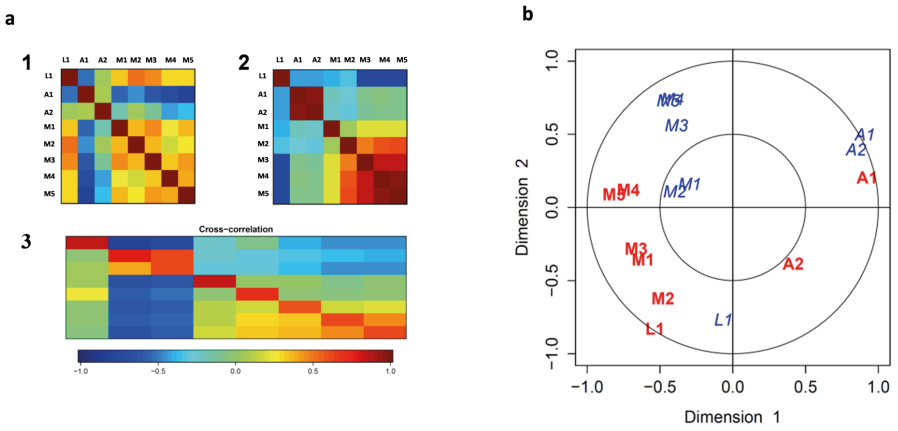
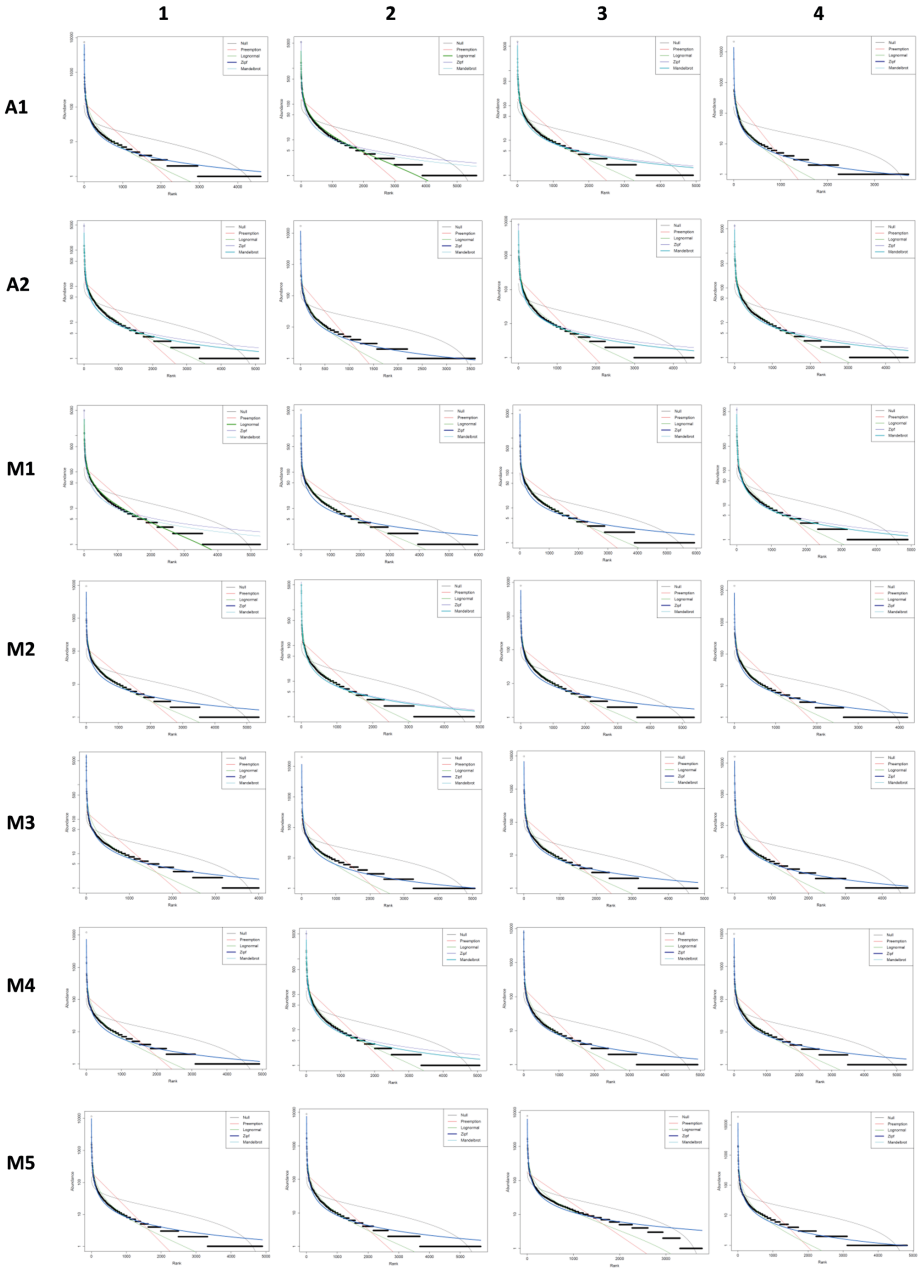


Fig. S7. Correspondence between similarity matrices of bean genotypic data and rhizobacterial families. (A) Correlation matrices retrieved from the rcc command (CCA package) for genotypic matrix (1), 16S microbiome matrix (2) and cross-correlation of both matrices (3). Blue indicates a negative association, green indicates null association and red indicates a positive association. (B) Bidimensional plot of Regularized Canonical Correspondence Analysis. Codes in blue were computed from bean DArT genotypic data and codes in red were computed from 16S microbiome data.


1
2
3
4
A1
A2
M1
M2
M3
M4
M5

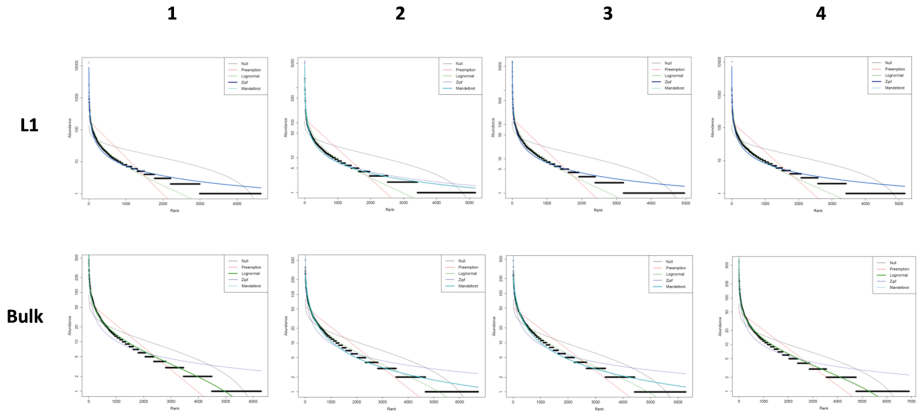


Fig. S8. **Species abundance distributions (SAD) models.** Broken stick (Null), Geometric (pre-emption), lognormal, Zipf and Zipf-Mandelbrot models were calculated using the bacterial read counts from the rhizosphere of the eight bean accessions as well as from the bulk soil. The legend indicates the bean accession to which the four plots in a same row belong.

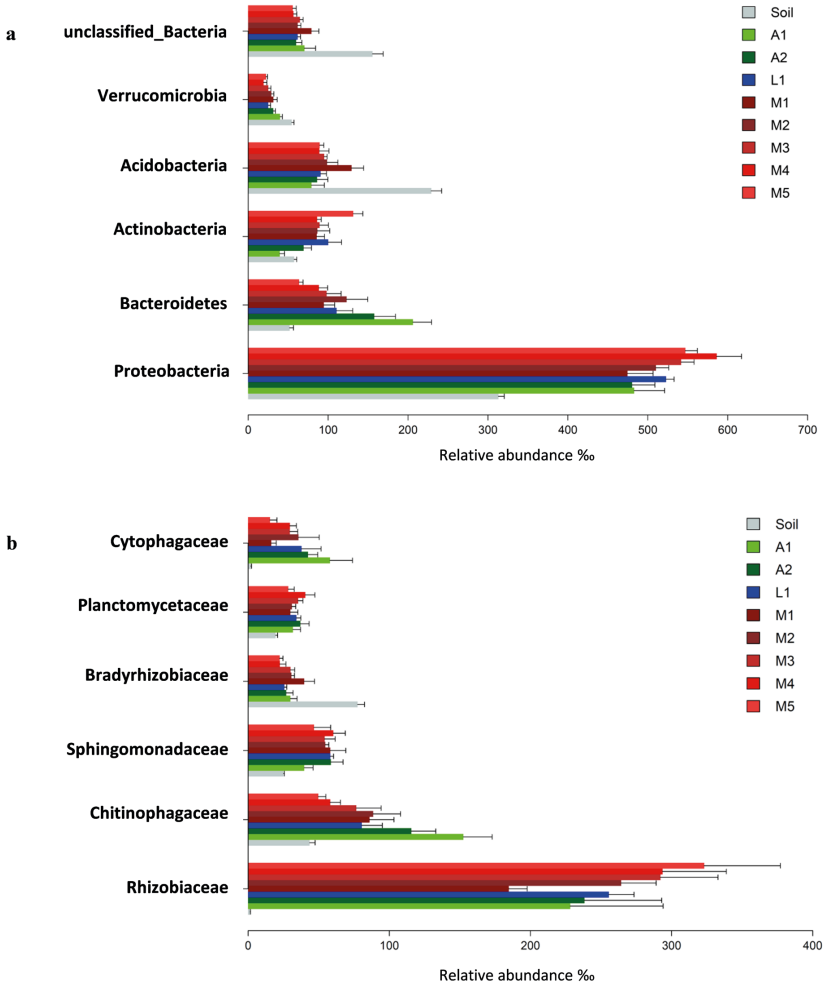


Fig. S9. Rhizosphere bacterial community structure of the bean accessions. Average relative abundance (RA ± SEM) of four replicates per accession of the most abundant bacterial Phyla **(a)** and **(b)** Families (taxonomically assigned) in the rhizosphere of the eight bean accessions and in bulk soil.

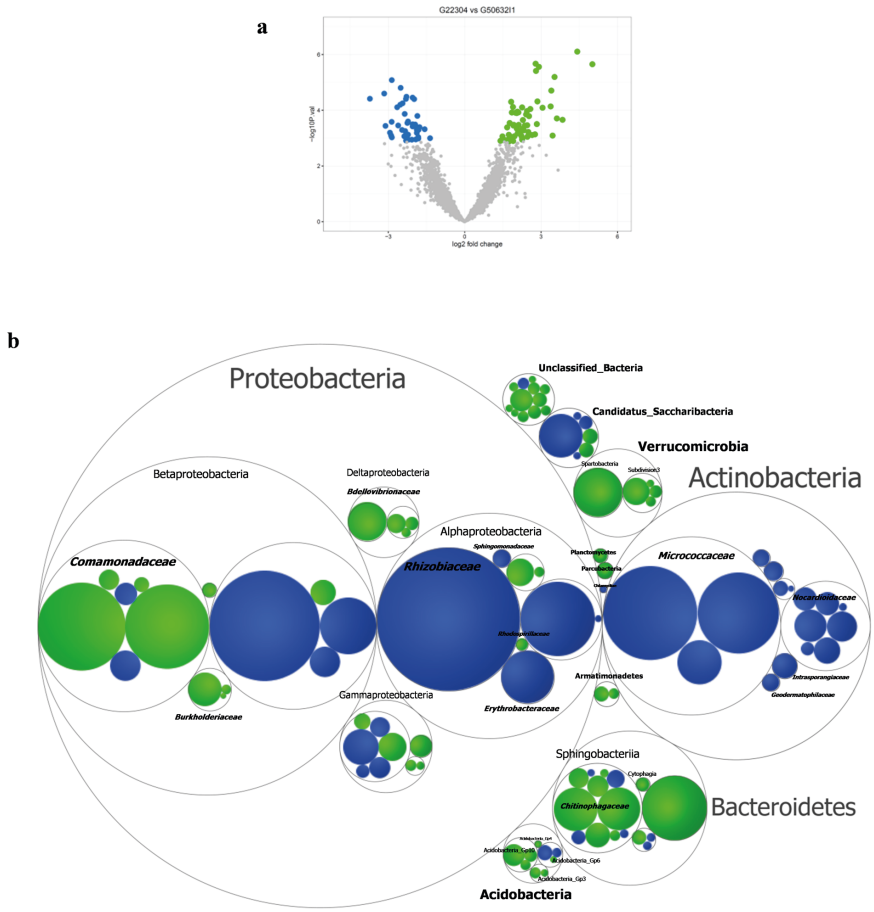
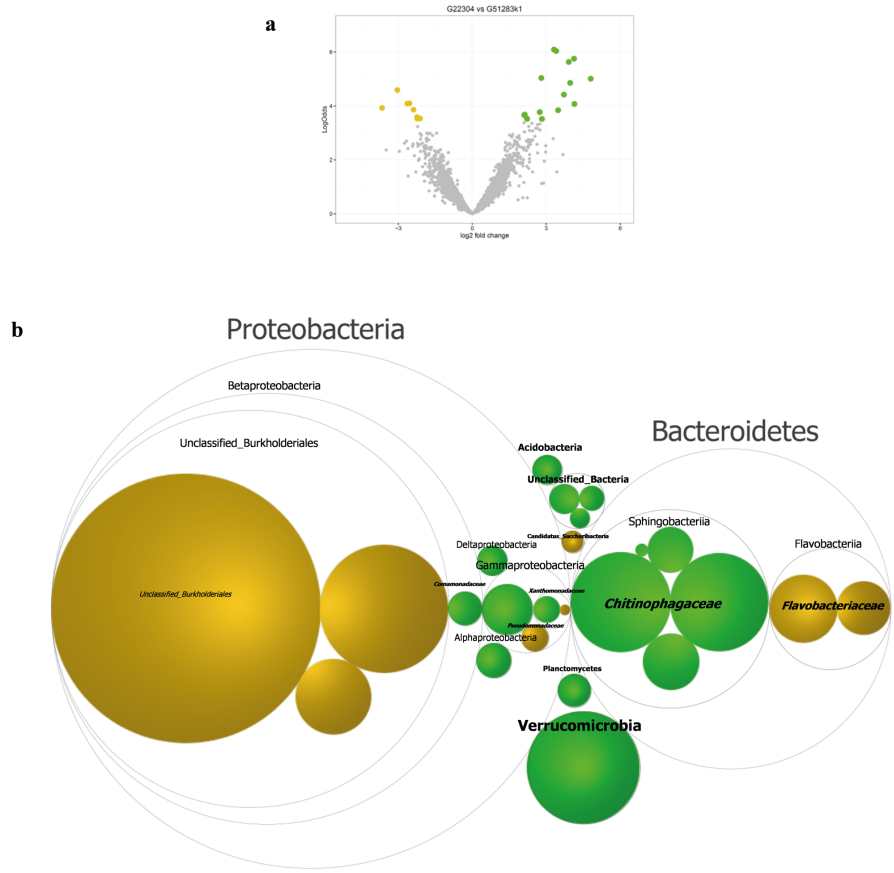


Fig. S10. Differential abundance of rhizobacterial OTUs between bean accessions A1 and L1. (a). Volcano plot showing the differentially abundant OTUs between wild accession A1 (green) and landrace accession L1 (blue). The plot was constructed using the significance and the log fold change of the comparison of the OTU abundances using a zero-inflated Gaussian distribution mixture model followed by an empirical Bayesian model. The y-axis represents the -log10P.val, which shows the significance of the differential abundance; the x-axis represents the log2 of the fold change for each OTU. Colorful dots are differentially abundant OTUs with an adjusted P-value (BH) < 0.05. (b). Bubble plot showing the differentially abundant OTUs between accessions A1 and L1. For this, the same data as for the volcano plot is used with the taxonomy assigned to each OTU. The biggest circles represent Phylum level, the inner circles Class and Family levels. The colored circles represent OTUs enriched in the rhizosphere of accession A1 (green) or enriched in the rhizosphere of accession L1 (blue). The size of the circle represents the mean read abundance of the differentially abundant OTU.



Supplementary Figure S11. Differential abundance of rhizobacterial OTUs between bean accessions A1 and M1. (a). Volcano plot showing the differentially abundant OTUs between wild accession A1 (green) and landrace accession M1 (gold). The plot was constructed using the significance and the log fold change of the comparison of the OTU abundances using a zero-inflated Gaussian distribution mixture model followed by an empirical Bayesian model. The y-axis represents the -log10P.val, which shows the significance of the differential abundance; the x-axis represents the log2 of the fold change for each OTU. Colorful dots are differentially abundant OTUs with an adjusted P-value (BH) < 0.05. (b). Bubble plot showing the differentially abundant OTUs between accessions A1 and M1. For this, the same data as for the volcano plot is used with the taxonomy assigned to each OTU. The biggest circles represent Phylum level, the inner circles Class and Family levels. The colored circles represent OTUs enriched in the rhizosphere of accession A1 (green) or enriched in the rhizosphere of accession M1 (gold). The size of the circle represents the mean read abundance of the differentially abundant OTU.

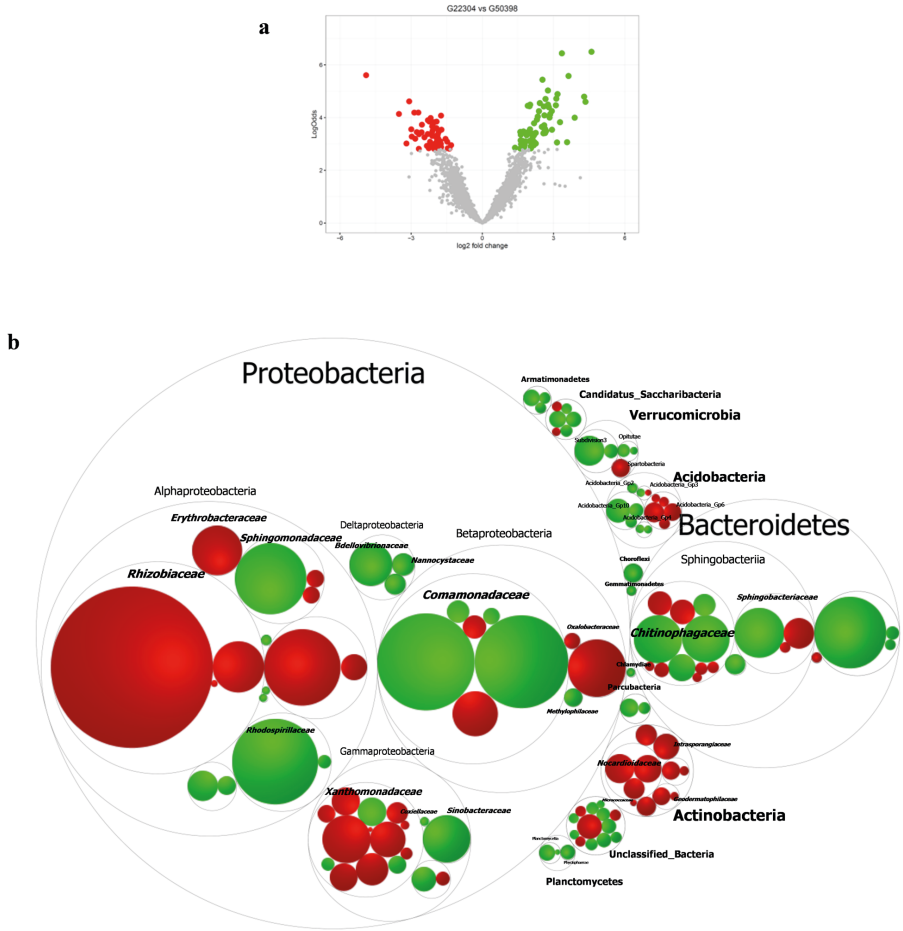


Fig. S12. Differential abundance of rhizobacterial OTUs between bean accessions A1 and M2. (a). Volcano plot showing the differentially abundant OTUs between wild accession A1 (green) and landrace accession M2 (red). The plot was constructed using the significance and the log fold change of the comparison of the OTU abundances using a zero-inflated Gaussian distribution mixture model followed by an empirical Bayesian model. The y-axis represents the -log10P.val, which shows the significance of the differential abundance; the x-axis represents the log2 of the fold change for each OTU. Colorful dots are differentially abundant OTUs with an adjusted P-value (BH) < 0.05. (b). Bubble plot showing the differentially abundant OTUs between accessions A1 and M2. For this, the same data as for the volcano plot is used with the taxonomy assigned to each OTU. The biggest circles represent Phylum level, the inner circles Class and Family levels. The colored circles represent OTUs enriched in the rhizosphere of accession A1 (green) or enriched in the rhizosphere of accession M2 (red). The size of the circle represents the mean read abundance of the differentially abundant OTU.

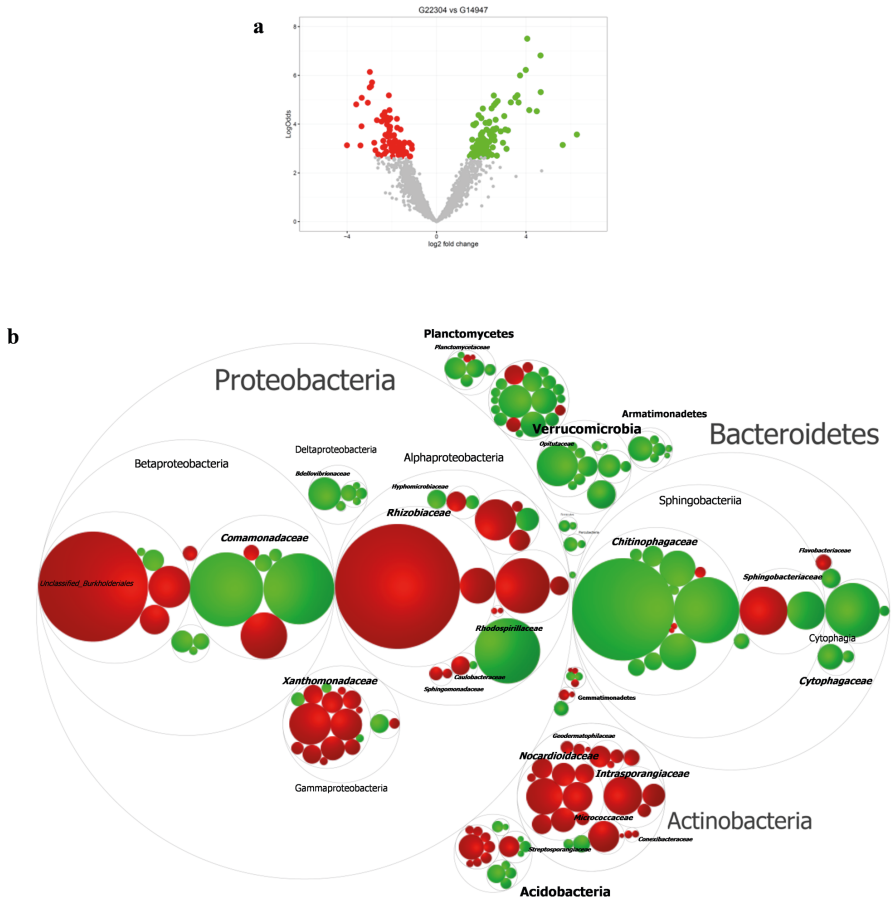


Fig. S13. Differential abundance of rhizobacterial OTUs between bean accessions A1 and M3. (a). Volcano plot showing the differentially abundant OTUs between wild accession A1 (green) and landrace accession M3 (red). The plot was constructed using the significance and the log fold change of the comparison of the OTU abundances using a zero-inflated Gaussian distribution mixture model followed by an empirical Bayesian model. The y-axis represents the -log10P.val, which shows the significance of the differential abundance; the x-axis represents the log2 of the fold change for each OTU. Colorful dots are differentially abundant OTUs with an adjusted P-value (BH) < 0.05. (b). Bubble plot showing the differentially abundant OTUs between accessions A1 and M3. For this, the same data as for the volcano plot is used with the taxonomy assigned to each OTU. The biggest circles represent Phylum level, the inner circles Class and Family levels. The colored circles represent OTUs enriched in the rhizosphere of accession A1 (green) or enriched in the rhizosphere of accession M3 (red). The size of the circle represents the mean read abundance of the differentially abundant OTU.

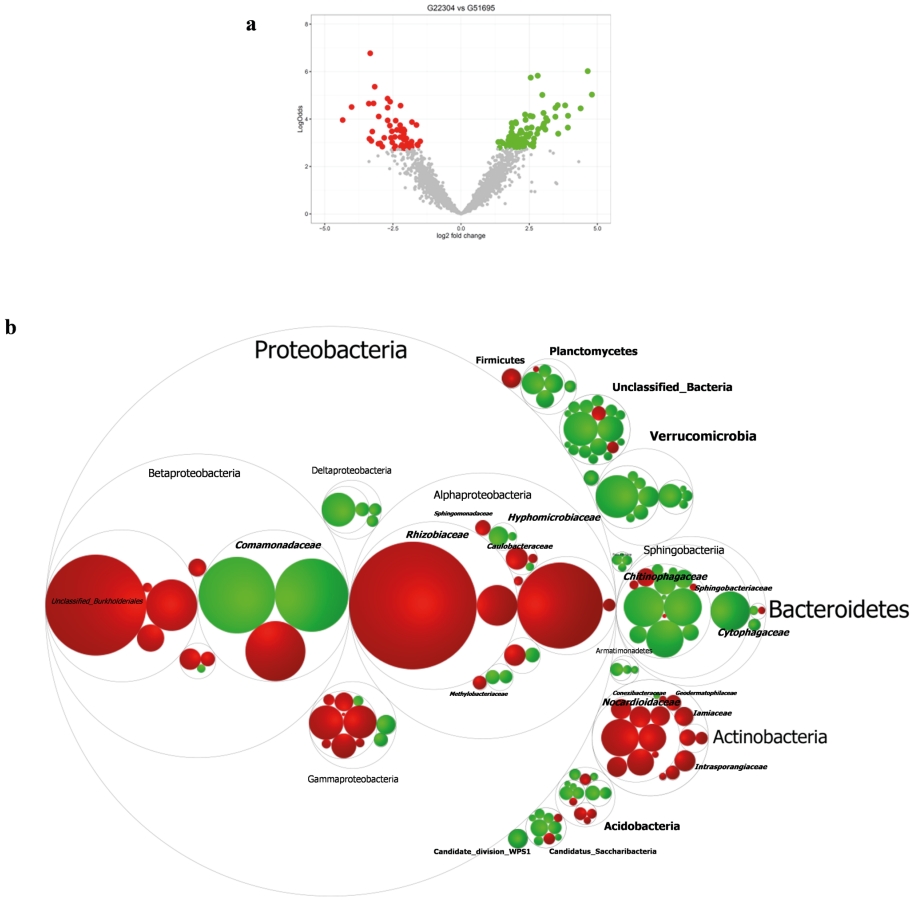


Fig. S14. Differential abundance of rhizobacterial OTUs between bean accessions A1 and M4. (a). Volcano plot showing the differentially abundant OTUs between wild accession A1 (green) and landrace accession M4 (red). The plot was constructed using the significance and the log fold change of the comparison of the OTU abundances using a zero-inflated Gaussian distribution mixture model followed by an empirical Bayesian model. The y-axis represents the -log10P.val, which shows the significance of the differential abundance; the x-axis represents the log2 of the fold change for each OTU. Colorful dots are differentially abundant OTUs with an adjusted P-value (BH) < 0.05. (b). Bubble plot showing the differentially abundant OTUs between accessions A1 and M4. For this, the same data as for the volcano plot is used with the taxonomy assigned to each OTU. The biggest circles represent Phylum level, the inner circles Class and Family levels. The colored circles represent OTUs enriched in the rhizosphere of accession A1 (green) or enriched in the rhizosphere of accession M4 (red). The size of the circle represents the mean read abundance of the differentially abundant OTU.

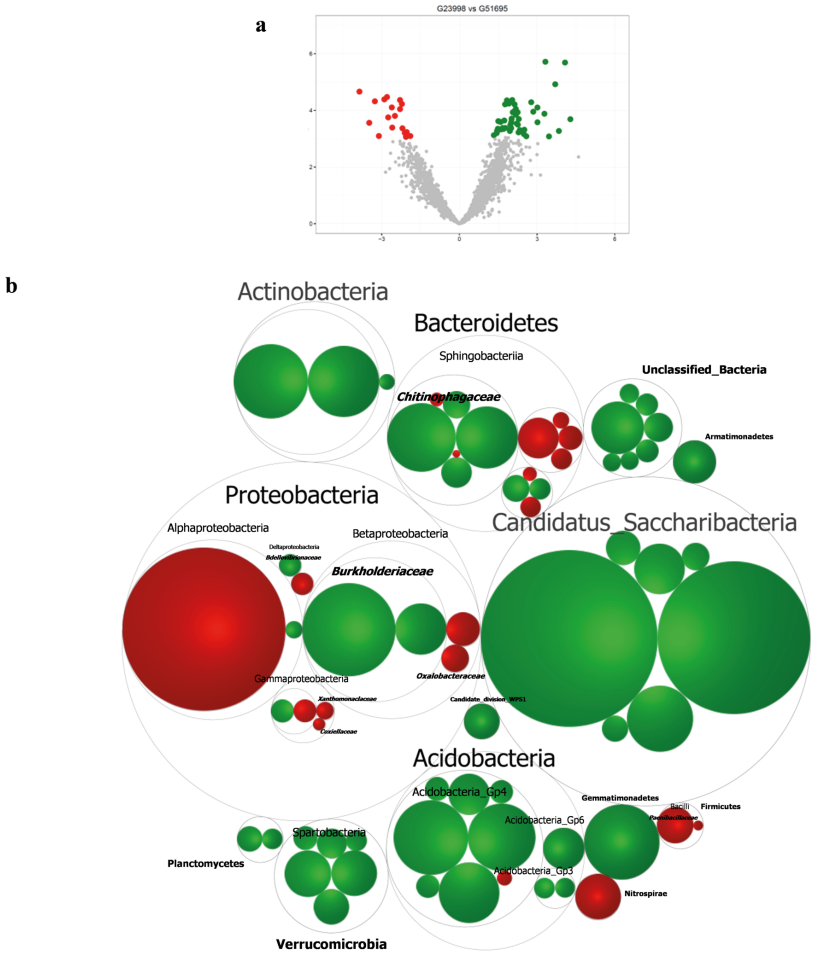


Fig. S15. Differential abundance of rhizobacterial OTUs between bean accessions A2 and M4. (a). Volcano plot showing the differentially abundant OTUs between wild accession A2 (green) and landrace accession M4 (red). The plot was constructed using the significance and the log fold change of the comparison of the OTU abundances using a zero-inflated Gaussian distribution mixture model followed by an empirical Bayesian model. The y-axis represents the -log10P.val, which shows the significance of the differential abundance; the x-axis represents the log2 of the fold change for each OTU. Colorful dots are differentially abundant OTUs with an adjusted P-value (BH) < 0.05. (b). Bubble plot showing the differentially abundant OTUs between accessions A2 and M4. For this, the same data as for the volcano plot is used with the taxonomy assigned to each OTU. The biggest circles represent Phylum level, the inner circles Class and Family levels. The colored circles represent OTUs enriched in the rhizosphere of accession A2 (green) or enriched in the rhizosphere of accession M4 (red). The size of the circle represents the mean read abundance of the differentially abundant OTU.

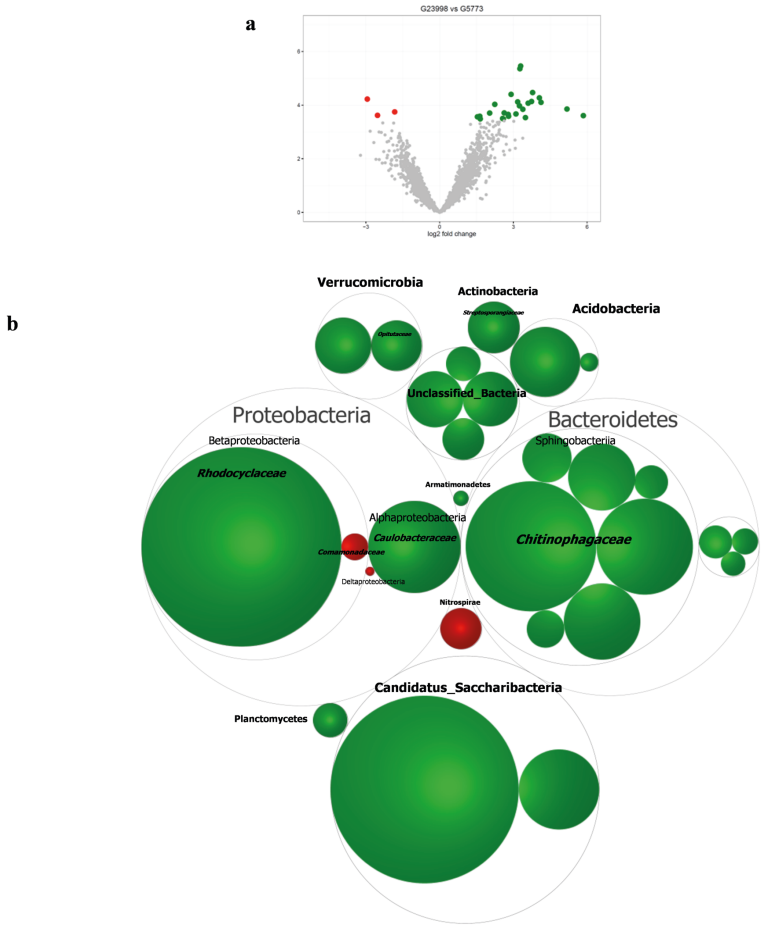


Fig. S16. Differential abundance of rhizobacterial OTUs between bean accessions A2 and M5. (a). Volcano plot showing the differentially abundant OTUs between wild accession A2 (green) and landrace accession M5 (red). The plot was constructed using the significance and the log fold change of the comparison of the OTU abundances using a zero-inflated Gaussian distribution mixture model followed by an empirical Bayesian model. The y-axis represents the -log10P.val, which shows the significance of the differential abundance; the x-axis represents the log2 of the fold change for each OTU. Colorful dots are differentially abundant OTUs with an adjusted P-value (BH) < 0.05. (b). Bubble plot showing the differentially abundant OTUs between accessions A2 and M5. For this, the same data as for the volcano plot is used with the taxonomy assigned to each OTU. The biggest circles represent Phylum level, the inner circles Class and Family levels. The colored circles represent OTUs enriched in the rhizosphere of accession A2 (green) or enriched in the rhizosphere of accession M5 (red). The size of the circle represents the mean read abundance of the differentially abundant OTU.

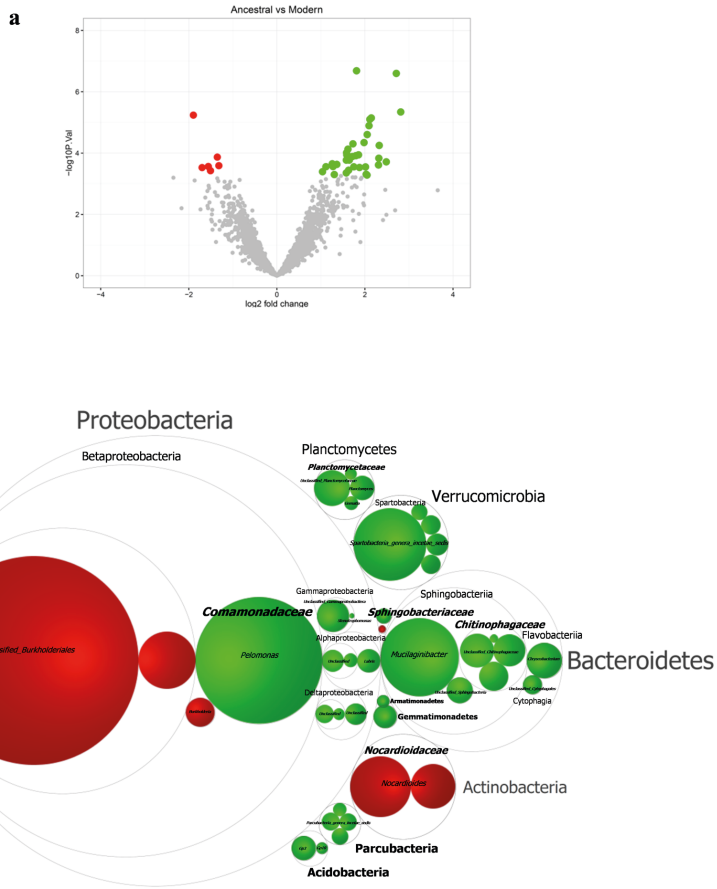


Fig. S17. Differential abundance of rhizobacterial OTUs between groups of bean accessions. (a). Volcano plot showing the differentially abundant OTUs between the wild accessions (green) vs modern accessions (red). The plot was constructed using the significance and the log fold change of the comparison of the OTU abundances using a zero-inflated Gaussian distribution mixture model followed by an empirical Bayesian model. The y-axis represents the -log10P.val, which shows the significance of the differential abundance; the x-axis represents the log2 of the fold change for each OTU. Colorful dots are differentially abundant OTUs with an adjusted P-value (BH) < 0.05. (b). Bubble plot showing the differentially abundant OTUs between ancestral accessions vs the modern accessions. For this, the same data as for the volcano plot is used with the taxonomy assigned to each OTU. The biggest circles represent Phylum level, the inner circles Class and Family levels. The colored circles represent OTUs enriched in the rhizosphere of ancestral accessions (green) or enriched in the rhizosphere of modern accessions (red). The size of the circle represents the mean read abundance of the differentially abundant OTU.

Table S1. Basic passport information of the common bean accessions used in this study.

Genetic pool	CIAT classification	DArT classification	Accession number	Weight of 100 seeds (g)	Seed Storage Protein*	Growth type	Country of origin	Department	Municipality	Altitude (m.a.s.l.)
Mesoamerican	Wild	A1	G22304	7.3	B, CH	Climbing	Colombia	Cundinamarca	Ubaqué	1750
Mesoamerican	Wild	M1	G51283K1	14.8	S, Sd, CH	Climbing	Colombia	Antioquia	Liborina	1930
Mesoamerican	Landrace	L1	G50632I1	24.7	CH	Climbing	Colombia	Antioquia	Andes	1750
Mesoamerican	Landrace	M2	G50398	22.8	S	Climbing	Colombia	Antioquia	Liborina	1970
Mesoamerican	Landrace	A2	G23998	23.6	B, CH	Climbing	Colombia	Cundinamarca	Tena	1540
Mesoamerican	Modern	M5	G5773	24	B	Bush	Colombia	Valle del Cauca	Palmira	800-1500
Mesoamerican	Modern	M3	G14947	19	S	Climbing	Colombia	Cundinamarca	Villeta	n.a.
Mesoamerican	Modern	M4	G51695	19.9	S	Climbing	Colombia	n.a.**	n.a.	n.a.
*CH, B, S and Sd are typical Mesoamerican phaseolin types. ** Not available information										

The plant material was kindly provided by the Genetic Resources Program of the International Center for Tropical Agriculture. Used with permission.

Table S2. Soil analysis. Physical and chemical characterization of the agricultural soil used to grow the common bean plants for the microbiome characterization and for the root morphology analysis.

Soil physicochemical analysis - Carmen soil		
Texture (%)	**Clay**	8
	Silt	30
	Sand	62
	Classification	Clay loam
dSm-1	**pH**	5.8
OM (%)		17.9
cmolc kg-1	**Al**	ND
	Ca	15.0
	Mg	2.9
	K	0.6
	Na	0.03
CEC		18.50
mg kg-1	**P**	56
	S	8
	Fe	50
	Mn	2
	Cu	3
	Zn	7
	B	0.24
Methods:		
Texture: Bouyoucos; pH: Water (1:1); Organic matter: Walkley Black; Al: KCl 1M; Ca, Mg, K, Na: Ammonium acetate 1M; CEC: Cation exchange sum; S: monocalcium phosphate 0.008M; Fe, Mn, Cu, Zn: Olsen-EDTA; B: Hot water; P: Bray II.		
Units consideration:		
ND: Not detectable; dSm-1 = mmho cm-1; cmolc kg-1 = meq/100 g soil; mg kg-1 = ppm		

Analysis performed in the Laboratory of Soils at the National University of Colombia – Medellin.

Table S3. Root traits of the 8 bean accessions. Average values (±SEM) of three replicates per accession.

	Common bean accessions							
	A1	**A2**	**L1**	**M1**	**M2**	**M3**	**M4**	**M5**
Length(cm)	313.4±74.7	344.7±4.1	527.3±27.0	169.6±17.1	336.3±34.5	323.7±124.1	375.6±40.8	258.6±79.4
SurfArea(cm^2)	42.8±15.4	50.4±7.3	64.5±4.3	36.4±2.8	42.3±5.0	41.8±21.7	44.1±6.1	40.8±6.8
AvgDiam(mm)	0.414±0.051	0.467±0.073	0.389±0.008	0.697±0.084	0.402±0.037	0.378±0.063	0.372±0.024	0.589±0.142
RootVolume(cm^3)	0.480±0.227	0.616±0.179	0.628±0.052	0.640±0.105	0.430±0.083	0.454±0.298	0.415±0.078	0.557±0.053
Tips	1075.3±176.6	1151.0±260.7	2437.0±73.9	375.5±82.9	1502.0±387.9	1200.0±458.1	1662.7±131.8	905.7±377.3
Forks	842.0±327.1	805.3±108.7	1352.0±164.5	420.3±6.6	660.3±41.6	716.3±393.2	728.0±124.2	516.0±162.8
Crossings	167.3±76.6	135.7±15.9	222.3±47.5	68.3±8.1	105.7±21.9	126.7±71.1	122.7±25.4	73.7±34.6
0 to 0.2mm diameter	68.2±22.9	180.4±32.5	184.4±76	35.3±3.9	202.7±32.7	190.4±64	239.1±24.2	130.5±58.5

All 8 accessions were grown in agricultural soil until growth stage V4 was reached. Entire root systems were sampled, stained, scanned and analyzed with WinRHIZO software.

Table S4. 16S rRNA Sequencing data obtained in this study. The sample ID, the number of base pairs, number of reads and the average length of the amplicon per sample are shown.

Sample_ID	Base pairs	Number of reads	Avg Length
A1_1	26595518	64997	409.2
A1_2	34345866	84122	408.3
A1_3	30665763	75428	406.6
A1_4	32453285	79203	409.7
A2_1	29748983	72959	407.7
A2_2	30132023	73955	407.4
A2_3	26574830	65351	406.6
A2_4	33881173	83180	407.3
L1_1	31311478	76948	406.9
L1_2	31870563	78308	407
L1_3	27922849	68568	407.2
L1_4	29834453	73264	407.2
M1_1	32425257	79436	408.2
M1_2	27258923	66688	408.8
M1_3	29402164	72006	408.3
M1_4	28102272	68925	407.7
M2_1	29143734	71465	407.8
M2_2	27128614	66661	407
M2_3	29573226	72471	408.1
M2_4	27327531	67044	407.6
M3_1	31045652	76239	407.2
M3_2	26398887	65017	406
M3_3	31684990	77582	408.4
M3_4	31096034	76435	406.8
M4_1	30254083	74590	405.6
M4_2	33906760	83322	406.9
M4_3	32935953	80801	407.6
M4_4	32726126	80218	408
M5_1	32112192	79129	405.8
M5_2	26871775	65775	408.5
M5_3	32686819	80526	405.9
M5_4	27123268	66657	406.9
Bulk_1	25992115	63213	411.2
Bulk_2	26659442	64912	410.7
Bulk_3	23636980	57543	410.8
Bulk_4	25508498	62121	410.6

This data is obtained after quality filtering and merging of paired end reads (see methods)

Table S5. AIC values for six rank abundance distribution models to test niche neutral models. The lowest AIC (Akaike Information Criterion) value for each sample represents the best fit model which is shown in bold

	AIC					
Sample	**Broken-stick**	**Pre-Emption**	**Log-Normal**	**Zipf**	**Zipf-Mandelbrot**	**ZSM**
A1	136293.7	125529.8	25818.2	**16495.8**	16947.8	38977.0
A1	125448.2	104990.9	**23140.4**	26520.8	24756.2	52983.0
A1	141830.2	122271.6	22999.4	20010.7	**19525.3**	41938.8
A1	280048.0	239534.0	40971.0	**24824.0**	24826.0	35671.6
A2	117971.8	99879.0	21549.0	22224.2	**20176.8**	43350.4
A2	230522.6	202648.2	33845.4	**17429.3**	17431.3	33026.8
A2	164766.8	130430.7	21622.5	22258.8	**20894.7**	38277.2
A2	135060.3	113990.9	20920.3	18781.3	**17829.5**	41709.8
L1	189645.0	163219.0	29238.0	**21238.0**	21240.0	41137.4
L1	136843.1	123173.4	25467.9	**19173.8**	18071.4	43057.4
L1	152133.6	134525.6	25090.0	**17672.7**	17674.7	39893.8
L1	163933.3	146460.4	29931.2	**22608.9**	22610.9	43008.8
M1	114414.4	93495.0	**21255.0**	25695.2	23392.2	44533.0
M1	96247.6	92261.0	27336.1	**23568.0**	23570.0	49129.0
M1	115010.8	107690.9	28252.1	**23586.6**	23588.6	45559.6
M1	140208.2	121647.2	23569.8	19296.1	**17693.4**	43386.4
M2	142555.4	130355.0	30664.8	**24190.8**	24192.8	43164.8
M2	131399.9	118383.5	24660.7	**17710.0**	17410.1	40616.2
M2	130914.5	119046.7	27736.7	**22995.5**	22997.5	44153.4
M2	176980.0	158412.0	33484.0	**23047.0**	23049.0	38716.2
M3	166981.8	158440.7	36006.3	**21761.2**	21763.2	39556.4
M3	133450.1	110675.2	**21376.5**	21889.1	19361.0	42384.6
M3	185977.1	163999.9	28010.3	**17252.1**	17254.1	38922.2
M3	167413.2	153616.8	31301.4	**20707.5**	20709.5	39195.8
M4	209316.5	180257.7	29408.9	**20394.5**	20396.5	40954.4
M4	171682.8	154887.2	30450.3	**21157.5**	21159.5	44934.8
M4	119442.1	118007.4	25516.4	**18994.7**	18996.7	38647.4
M4	255131.9	236237.6	44710.0	**23344.5**	23346.5	40673.6
M5	148212.4	140938.7	28499.0	**15828.5**	15830.5	37933.2
M5	267760.0	252712.0	53232.0	**28932.0**	28934.0	42485.8
M5	148620.2	135906.7	29416.9	**21037.3**	21039.3	41446.2
M5	243582.4	217176.3	35530.8	**18474.0**	18476.0	37446.6

The ZSM (zero-sum multinomial) was calculated with the TeTame software, which tests whether the data fits a null or neutral model. The other five models were retrieved from the vegan package in R, using the command radfit, which fits all other models.

Table S6. Degree of immigration rate into local communities from a regional pool.

Sample	m*
A1	0.28145
A1	0.183391
A1	0.265016
A1	0.0867219
A2	0.350056
A2	0.122949
A2	0.300635
A2	0.211872
L1	0.186452
L1	0.323108
L1	0.372738
L1	0.335062
M1	0.264119
M1	0.393276
M1	0.546438
M1	0.168541
M1	0.42198
M2	0.271972
M2	0.362148
M2	0.136219
G14947_1	0.371529
G14947_2	0.298588
G14947_3	0.377907
G14947_4	0.602484
G51695_1	0.280479
G51695_2	0.440871
G51695_3	0.125627
G51695_4	0.265491
G5773_1	0.168336
G5773_2	0.282096
G5773_3	0.282896
G5773_4	0.353746
Bulk_soil_1	**0.557585**
Bulk_soil_2	**0.784444**
Bulk_soil_3	**0.507921**
Bulk_soil_4	**0.867186**

*The *m* value was retrieved from a zero-sum multinomial distribution.
Values closer to 1 indicate no dispersal limitation.

Table S7. Differential abundance analysis of rhizobacteria on the eight bean accessions.

Accession	**A1**	**A2**	**L1**	**M1**	**M2**	**M3**	**M4**
A1							
A2	92						
L1	104	1					
M1	24	1	0				
M2	123	4	1	3			
M3	177	0	0	2	2		
M4	142	68	0	3	2	0	
M5	402	28	0	2	4	0	1

The filtered OTU table was fitted into a Zero Inflated Gaussian model (ZIG), and moderated t tests followed by an empirical Bayes approach were performed. *P* values were adjusted (BH) for false discovery rate (FDR < 0.05). In the table the number of differentially abundant OTUs in pair comparisons between bean accessions is shown.

Chapter 4

Deciphering the microbiome assembly of wild and modern common bean (*Phaseolus vulgaris*) grown in native and agricultural soils from Colombia

Juan E. Pérez-Jaramillo, Mattias de Hollander, Camilo A. Ramírez,

Rodrigo Mendes, Jos M. Raaijmakers and Víctor J. Carrión

Submitted for publication

Abstract

Modern crops are typically cultivated in agriculturally well-managed soils far from the centres of origin where their wild relatives thrive. How this physical decoupling impacted plant microbiome assembly is not well understood. Here, we investigated if the transition from a native to an agricultural soil affected rhizobacterial community assembly of wild and domesticated common bean (*Phaseolus vulgaris*) and if this led to a loss of rhizobacterial diversity. The impact of the bean genotype on rhizobacterial assembly was more prominent in the agricultural soil than in the native soil. Only 113 operational taxonomic units (OTUs) out of a total of 15,925 were shared by all eight bean accessions in both soils, representing 25.9% of all sequence reads. More OTUs were exclusively found in the rhizosphere of common bean in the agricultural soil as compared to the native soil, and in the rhizosphere of modern bean accessions as compared to wild accessions. Co-occurrence analyses further showed a reduction in complexity of the interactions going from native to agricultural soil. Collectively, these results suggest that habitat expansion of common bean from its native soil environment to an agricultural context had an overall positive effect on rhizobacterial diversity and led to a stronger bean genotype-dependent effect on rhizosphere microbiome assembly.

Keywords: native soil, agricultural soil, core microbiome, networks

Introduction

Plant domestication and the agricultural revolution provided a more continuous food supply to early human hunter-gatherer groups and were key drivers of the conformation of stable human settlements (Purugganan and Fuller 2009). Domestication led to major changes both in phenotypic and genotypic traits of crop varieties including larger seed size, loss of dispersal mechanisms and determinate growth (Gepts 2004; Doebley *et al*., 2006). However, domestication also led to a reduction in genetic diversity, referred to as the domestication syndrome (Doebley *et al*., 2006). Recent studies further showed that domestication affected rhizosphere microbiome assembly of several plant species, including sugar beet (Zachow *et al*., 2014), barley (Bulgarelli *et al*., 2015), sunflower (Leff *et al*., 2017), and common bean (Pérez-Jaramillo *et al*., 2017). For common bean, we previously revealed that wild accessions were enriched in bacterial taxa from the phylum Bacteroidetes, whilst modern accessions were enriched in Actinobacteria and that this compositional shift was associated with plant genotypic as well as root phenotypic traits (Pérez-Jaramillo *et al*., 2017).

Plant domestication not only comes with changes in plant traits, but is also accompanied by progressive changes in the habitat and crop management practices to promote high yields and to protect the domesticated plants from biotic and abiotic stress factors (Pérez-Jaramillo *et al*., 2016). Hence, the transition from native habitats to agricultural soils may have led to a loss of plant-associated microbes thereby affecting specific, co-evolved beneficial functions of the plant microbiome. For example, long-term nitrogen fertilization resulted in the evolution of less-mutualistic rhizobia, providing fewer benefits to the host (Weese *et al*., 2015). Similarly, it was shown that nitrogen amendments suppressed soil respiration and microbial biomass, promoting copiotrophs such as Actinobacteria and Firmicutes while reducing the abundance of oligotrophs such as Acidobacteria and Verrucomicrobia (Ramirez *et al*., 2012). It has been also shown that

the occurrence of members of the phylum Bacteroidetes, whose members are known for their abilities to degrade complex polymeric organic matter, is negatively affected by agricultural soil management practices (Wolińska *et al*., 2017). Similarly, conversion of the Amazon rainforest to agriculture resulted in biotic homogenization of soil bacterial communities and a net loss of microbial diversity (Rodrigues *et al*., 2013). For most crop plants, however, there is little knowledge on the co-evolutionary trajectory between plants and their microbiomes during habitat expansion and if domestication indeed led to a reduced microbial diversity and a loss of specific microbial genera in these new habitats. In this study, we used *Phaseolus vulgaris* (common bean) as 'model' plant species. Common bean originated in central Mexico and as a wild species spread throughout Central and South America (Gepts 1998; Bitocchi *et al*., 2012; Desiderio *et al*., 2013). Geographical isolation of wild common bean resulted in the establishment of the Mesoamerican and Andean genetic pools (Gepts and Bliss 1985) which were the basis of two independent domestication processes. As a consequence, domesticated common bean underwent several morphological and physiological changes as well as a significant reduction in genetic diversity (Gepts and Debouck 1991; Sonnante *et al*., 1994; Chacón *et al*., 2005). We selected wild, landraces and modern accessions of Mesoamerican common bean originating from Colombia based on a number of traits (Pérez-Jaramillo *et al*., 2017). Here, we hypothesized that the transition from a native soil environment to an agriculturally managed soil led to a loss of bacterial diversity in the plant rhizosphere microbiome. Therefore, the eight bean accessions were grown in a native soil and in an agricultural soil, both collected in the Colombian highlands, followed by rhizobacterial community profiling, species abundance modelling and co-occurrence network analyses.

Material and Methods

Selection of soils and plant accessions

Two types of soil were selected for this study in the north-west region of Colombia. The native soil was collected in a rural area near to the municipality of Angostura (Antioquia, 6° 53′ 7″ N, 75° 20′ 7″ W). This region has the same climatic conditions, altitude and local plant diversity that have been reported for wild common bean populations collected in Colombia (Toro *et al*., 1990). A successional forest was identified in the region and soil samples were taken from the top layer (0-20cm) after cleaning the litter, wood and unwanted material. Several landraces typically associated with Mesoamerican traits were collected near this region; therefore we refer to this soil as "native". The agricultural soil was collected in a common bean producing farm in rural area of the municipality of El Carmen de Viboral (6°4′55″ N, 75º20′3″W). This soil was under cultivation for the last 30 years in a crop rotation system composed of maize, common bean and potato. Tillage, liming, chemical fertilization (N-P-K) and the application of poultry waste are typical agricultural practices in the region. The climatic conditions, the altitude and the local plant diversity in this region are not suitable for the growth of Colombian Mesoamerican wild common bean populations, but are optimal for Andean domesticated common bean varieties. Physicochemical analyses were performed in the Soil Science Laboratory from the National University of Colombia in Medellín, using standard procedures (Table S1). Two wild, three landraces and three improved varieties (cultivars) of common bean (*Phaseolus vulgaris*) were selected according to the following characteristics: belong to the Colombian Mesoamerican genetic pool, same race, similar phaseolin type, same altitudinal range, adapted to the same climatic conditions and same growth type. The seeds were kindly provided by the Genetic Resources Program at the International Centre for Tropical Agriculture - CIAT – in Palmira, Colombia. A genotypic analysis was performed on the common bean accessions to validate the domestication status provided

by the original passport (Pérez-Jaramillo *et al*., 2017). As a result, we reclassified the accessions as two wild (A1, A2), one landrace (L1), and five modern accessions (M1 to M5).

Experimental Design

Seeds of the eight bean accessions were surface-sterilized twice with sodium hypochlorite (0.5%) during three minutes and rinsed in sterile water four times. 100 µL of the last rinsing step was cultured in Tryptic Soy Agar (TSA, Oxoid) and in Potato Dextrose Agar (PDA, Difco) media by triplicate in order to check the growth of bacteria and fungi, respectively. Disinfected seeds were germinated on filter paper with sterile tap water; after two to five days all the seeds had germinated. The native and agricultural soils were air dried, passed through a 2-mm-mesh sieve and distributed into 1 L PVC pots, with 700 g of dried soil per pot. Seedlings were transferred to the pots, with one plant per pot and four replicates per bean accession and per soil. The plants were cultivated in a growth chamber for one week and then arranged randomly in a screenhouse with an average temperature of 25ºC, 12h of daylight, and daily watered with tap water. Four pots with native soil and four pots with agricultural soil, both without plants were used as bulk soils.

Sampling of rhizospheric soil

At flowering stage, the plants were carefully removed from the pots keeping the root system intact. Soil loosely adhered to the roots was removed by vigorous shaking, and when no more soil could be removed, the root system was submerged in tubes with 5mL of LifeGuard Soil Preservation Solution (Mo Bio Laboratories, Carlsbad, CA, USA) and vigorously shaken in order to wash the roots and recover around 1g of rhizospheric soil per sample for total DNA isolation. For the bulk soils, approximately 1 g of soil was collected from each control pot and also submerged in 5 mL of the LifeGuard solution.

The LifeGuard Soil Preservation Solution can prevent microbial growth while maintaining nucleic acid integrity. All samples were kept at -20ºC until further use.

DNA isolation

For each plant accession in each soil, four replicates of rhizospheric soil were used for total DNA extraction as well as four replicates of control soil. To obtain the total DNA, a two-step approach was followed in order to recover RNA and DNA from the same sample. First, RNA was isolated using the RNA PowerSoil Isolation Kit (Mo Bio Laboratories, Carlsbad, CA, USA) according to the manufacturer's instructions, with slight modifications as follows. After adding the phenol:chloroform:isoamyl alcohol solution to the bead tube containing the bead solution, solution SR1 and SR2, as well as the soil sample, the agitation step was applied for 40 min. This modification allowed us to increase the RNA yields. The RNA was stored at -80ºC for further use. For DNA isolation, the RNA PowerSoil® DNA Elution Accessory Kit (Mo Bio Laboratories, Carlsbad, CA, USA) was used. Briefly, after elution of the RNA from the RNA capture column, this column was transferred to another tube and the DNA elution procedure was performed according to manufacturer's instructions. Each obtained DNA sample was then cleaned with the PowerClean® DNA Clean-Up Kit (Mo Bio Laboratories, Carlsbad, CA, USA). Agarose gel electrophoresis and a ND1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) were used to control RNA and DNA yield and quality. DNA samples were stored at -80° C until further use.

16S amplicon sequencing and preparation of the OTU table

Total DNA was used to amplify and sequence the V3-V4 region of the 16S rRNA that creates an amplicon of approximately ~460bp. The library preparation and the sequencing protocol were done according to the general guidelines of Illumina MiSeq Reagent Kit

v3 (Illumina 2013). The RDP extension to PANDASeq (Masella *et al*., 2012) named Assembler (Cole *et al*., 2014) was used to merge paired-end reads with a minimum overlap of 10 bp and at least a PHRED score of 25. Primer sequences were removed from the per sample FASTQ files using Flexbar version 2.5 (Dodt *et al*., 2012). Sequences were converted to FASTA format and concatenated into a single file. All reads were clustered into OTUs using the UPARSE strategy by dereplication, sorting by abundance with at least two sequences and clustering using the UCLUST smallmem algorithm (Edgar, 2010). These steps were performed with VSEARCH version 1.0.10 (Rognes *et al*., 2015), which is an open-source and 64-bit multithreaded compatible alternative to USEARCH. Next, chimeric sequences were detected using the UCHIME algorithm (Edgar *et al*., 2011) implemented in VSEARCH. All reads before the dereplication step were mapped to OTUs using the usearch_global method implemented in VSEARCH to create an otu table and converted to BIOM-Format 1.3.1 (McDonald *et al*., 2012). Finally, taxonomic information for each OTU was added to the BIOM file by using the RDP Classifier version 2.10 (Cole *et al*., 2014). All steps were implemented in a Snakemake workflow (Köster and Rahmann 2012). The sequence data are deposited at the European Nucleotide Archive (ENA) under accession number PRJEB26084.

Diversity and abundance analysis

For downstream analysis we took the obtained OTU table and prepared a "filtered table" using QIIME (1.9.1) custom scripts (Kuczynski *et al*., 2012). First, we extracted from the OTU table the Bacteria domain using the command *split_otu_table_by_taxonomy.py*. Next, we discarded singletons, doubletons, Chloroplast and Mitochondria sequences using the command *filter_otus_from_otu_table.py*. With the "filtered_OTUtable", we first calculated the alpha diversity. Using the command *alpha_rarefaction.py*, the OTU table was rarefied to counts up to 35,000 reads. The reason to use this value was because

this was the lowest sequencing depth obtained from a sample. To calculate the diversity indexes, we used the *alpha_diversity.py* and *alpha_rarefaction* commands, obtaining Shannon, Observed OTUs, Chao1 and Faith's Phylogenetic Diversity metrics. One-way ANOVA and Tukey HSD, as well as statistical tests to validate ANOVA assumptions were performed in R (3.4.1) (R Core Team, 2015). For the Beta diversity, the unrarefied "filtered_OTUtable" was first normalized using the R package metagenomeSeq (v.1.12) (Paulson *et al*., 2016). We used a cumulative-sum scaling (CSS) method to avoid the biases generated with current sequencing technologies due to uneven sequencing depth (Paulson *et al*., 2013). With the normalized OTU table we calculated Bray-Curtis and weighted and unweighted Unifrac dissimilarity matrices and use it to build Principal Coordinate with Phyloseq package (v.1.10) (McMurdie and Holmes, 2013). The nonparametric *adonis* test was used to assess the percentage of variation explained by the soil type along with its statistical significance using Vegan (v.2.4-0) package (Oksanen *et al*. 2016), all performed in R. For the differential abundance analysis and the construction of the heat maps the STAMP software (v.2.1.3) was used (Parks *et al*., 2014). Rarefied OTU tables from Agricultural and Native rhizosphere and bulk soil data (35,000 reads per sample) were used for pairwise comparisons. Welch's t-tests followed by Bonferroni corrections were performed at Phylum and at Family level between soils. Dendograms were built in STAMP with an average neighbor method (UPGMA) and the rows included all the bacterial phyla observed in rhizosphere and bulk soil samples along with their relative abundance.

In order to compare the number of shared and exclusive rhizobacterial genera between common bean accessions in agricultural and native soil, we selected wild accessions A1 and A2 and modern bean accessions M3 and M4. These 4 accessions were selected in order to normalize the number of samples as well as the number of reads to compare. Regarding the modern accessions, M3 and M4 showed to be the most modern accessions

available both for native and agricultural soils (Pérez-Jaramillo *et al*., 2017). To depict the taxa exclusively found in a particular soil or accession we used the online tool Venny (2.1) (Oliveros, 2015) and to graphically represent the exclusive genera we built euler diagrams using the shiny app eulerr of the homonymous R package (Larsson, 2017). Euler diagrams are area-proportional generalized venn diagrams for which the requirement that all intersections be present is relaxed. Euler diagrams were built for the exclusive and shared genera per soil type and plant domestication status.

In order to have a better understanding of the composition of the bacterial diversity in the rhizosphere of the common bean accessions, we calculated several species abundance distribution models and determined whether neutral or niche-based mechanisms were governing the bacterial assembly. We hypothesized that the agricultural soil would be driven by neutral-based processes and that the native soil would respond to niche-based process. We used the command Radfit from the R package vegan to evaluate several abundance models and a zero-sum multinomial (ZSM) model. Species abundance distributions models and comparison of the models fit based on the Akaike Information Criterion (AIC) were calculated as previously reported (Pérez-Jaramillo *et al*., 2017).

Core microbiome and co-occurrence network analyses

For the core microbiome analyses, rarefied OTU tables (35,000 reads each sample) were used for both soils. The QIIME command *compute_core_microbiome.py* was applied in order to obtain a list of OTUs observed in all the common bean rhizosphere samples regardless of soil type. Core microbiome analyses were also performed for common bean on each soil type. Only core OTUs with a relative abundance > 0.5% were included for graphical purposes. Pie and donut charts were built in R. Network analysis was performed to assess the complexity of the interactions among microbial taxa in the common bean rhizosphere grown in the agricultural soil (n=32) and in the native soil (n=26). Best

practices for co-occurrence network construction were strictly followed. Rarefied OTU tables were filtered to a minimum threshold of 25 sequences per OTU. Non-random co-occurrence analyses were performed using SparCC (Friedman and Alm, 2012). *P*-values were obtained by 99 permutations of random selections of the data table. SparCC correlations with a magnitude > 0.8 or < -0.8 and statistically significant (P<0.01) were further included into network construction. The nodes in the reconstructed networks represent the OTUs at 97% identity, whereas the edges correspond to a strong and significant correlation between nodes. The topology of the network was inferred on a set of measures (number of nodes and edges, modularity, number of communities, average path length, network diameter, averaged degree and clustering coefficient) which were calculated using Gephi (v.0.9.2) (Bastian *et al*., 2009). Network visualizations were constructed using Cytoscape (v. 3.4.0) (Shannon *et al*., 2003). Clusters were calculated using a hierarchical clustering algorithm (HC-PIN) with the Cytoscape plugin Cytocluster (Li *et al*., 2017).

Results

Soil and genotype influence plant growth and development

Eight accessions of common bean, encompassing wild relatives, landraces and modern cultivars were grown in agricultural and native soils collected from the Colombian highlands, at the same time under the same screenhouse conditions (see Material and Methods). The soils differed in several physicochemical characteristics (Table 1). Briefly, the agricultural soil had more organic matter, available phosphorus and calcium as well as higher pH and cation exchange capacity (CEC) as compared to the native soil. The native soil showed higher concentrations of iron and aluminum. At flowering stage, each plant was harvested to collect rhizospheric soil and to assess several plant phenotypic traits. In the agricultural soil, significant differences were observed between the different

bean accessions in root dry weight and the number of days to reach flowering stage (Table 2). Genotype-dependent differences were also observed in the native soil (Table 2). Several replicates of the accession M5 did not grow in the native soil or showed a very poor development, and were therefore not included in further analyses. In general, the mean root dry weight was higher for bean accessions grown in the agricultural soil than in the native soil (Fig. S1a). Finally, the number of days to reach the flowering stage (R6) was higher in the native soil than in the agricultural soil, indicating a delayed development of the bean accessions in the native soil (Fig. S1b).

Table 1. Soil physical and chemical analyses. Physical and chemical characterization of the agricultural and the native soil used in this study to grow the common bean plants for the rhizobacterial characterization.

Unit	Item		Agricultural	Native
%	Texture	Clay	8	24
		Silt	30	12
		Sand	62	64
		Classification	Clay loam	Sandy clay loam
dSm-1	pH		5.8	4.7
%	Organic matter		17.9	11.6
cmolc kg-1	Al		ND	3.00
	Ca		15.0	0.3
	Mg		2.9	1.0
	K		0.6	0.4
	Na		0.03	0.03
	CEC		18.50	4.80
mg kg-1	P		56	3
	S		8	8
	Fe		50	620
	Mn		2	2
	Cu		3	3
	Zn		7	4
	B		0.24	0.21
Methods:				

Texture: Bouyoucos; pH: Water (1:1); Organic matter: Walkley Black; Al: KCl 1M; Ca, Mg, K, Na: Ammonium acetate 1M; CEC: Cation exchange sum; S: monocalcium phosphate 0.008M; Fe, Mn, Cu, Zn: Olsen-EDTA; B: Hot water; P: Bray II.
Units consideration:
ND: Not detectable; dSm-1 = mmho cm-1; cmolc kg-1 = meq/100 g soil; mg kg-1 = ppm

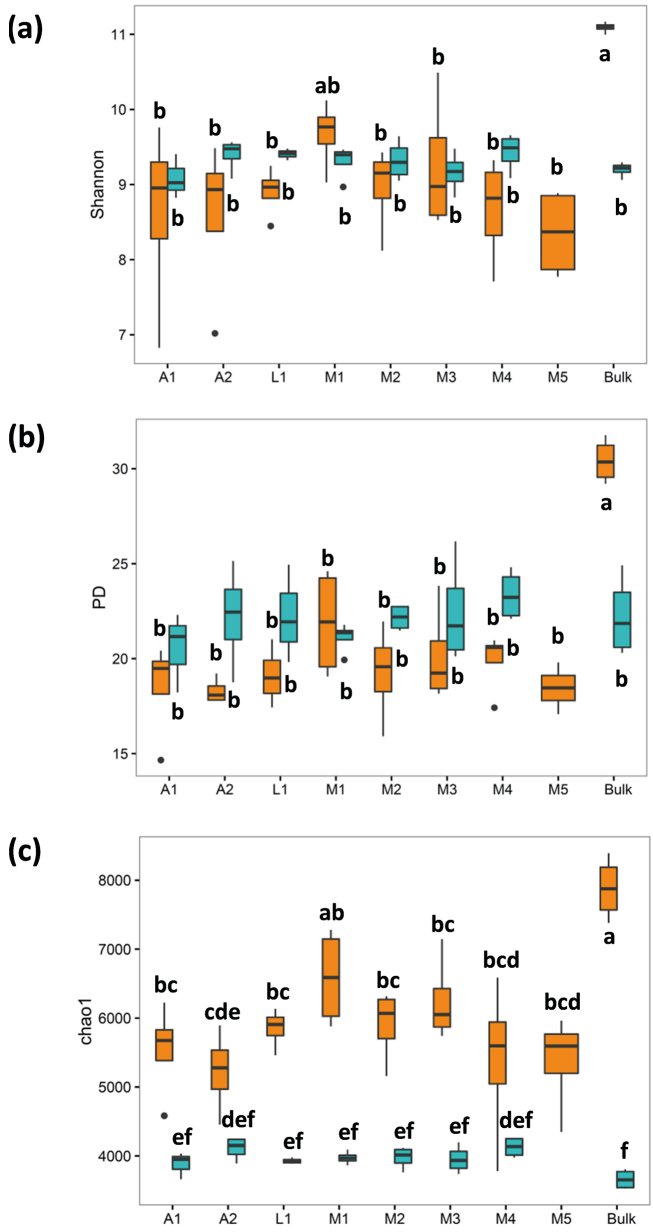


Fig. 1. Comparative analysis of the Alpha diversity of 16S rRNA rhizobacterial sequences from common bean accessions in agricultural and native soils. (a) Shannon, (b) Phylogenetic diversity, and (c) Chao 1 were calculated by soil type and for all bean accessions and the bulk soils. The data was rarefied up to 35,000 counts per sample. Statistically significant differences were determined by one-way ANOVA (P<0.05) followed by post hoc Tukey test. Cyan color was assigned to native soil and dark orange to agricultural samples.

Diversity of rhizobacterial communities is driven by soil type and rhizosphere

For bulk soil and rhizosphere samples, 4.24 million reads were recovered after quality filtering (Table S1), representing 16.727 bacterial operational taxonomic units (OTUs) at 97% sequence similarity. For the α-diversity, rarefaction curves were obtained for Chao1, observed OTUs and Phylogenetic Diversity (PD) indices (Fig. S2). It was observed that the evenness, represented by the Shannon index and the phylogenetic diversity (PD), were in general similar between rhizosphere samples in both soil types (Figs. 1a and b), while the bacterial species richness was significantly higher in the agricultural soil than in the native soil (Fig. 1c). In the agricultural soil, all diversity indices were significantly higher for the bulk soil than for the rhizosphere samples. Additionally, it was observed that agricultural bulk soil samples showed significantly higher values for all the diversity indices as compared to the native bulk soil samples. Regarding the β-diversity, Bray-Curtis metrics and Principal Coordinate Analysis (PCoA) revealed a significant effect of the soil type (Fig. 2a). Soil type alone explained 71.3% of the total variability in the bacterial community composition (PERMANOVA, $P<0.001$). Subsequently, the samples were divided by soil type and analyzed separately. For the total variability of the rhizobacterial community structure, the bean genotype explained 31.2% in the agricultural soil (PERMANOVA, $P<0.001$) and 28.3% in the native soil (PERMANOVA, $P<0.05$) (Figs. 2b and c). Unifrac metrics confirmed the results observed with Bray-Curtis metrics (Fig. S3). The mean Bray-Curtis distances showed that the variability of the rhizobacterial communities within samples of the same accession was significantly lower as compared to the distance between the bean accessions (Agricultural soil, $P<0.001$; Native soil, $P<0.05$; t-test, Bonferroni-corrected). For the agricultural soil, the rhizobacterial community composition of wild bean accession A1 was similar to that of wild accession A2, but significantly different from that of the landrace and the five modern bean accessions ($P<0.001$; t-test, Bonferroni-corrected). In

the native soil, however, rhizobacterial community composition did not differ significantly between the wild and modern bean accessions.

Table 2. Phenological traits of the common bean accessions grown in native and agricultural soils. Data for root dry weight and the number of days to reach the flowering stage is shown.

	Root dry weight (g)		Days to flower	
Accession	**Native**	**Agricultural**	**Native**	**Agricultural**
A1	0.082±0.015 bcd	0.438±0.141 ab	111.0±12.3 a	93.5±12.4 a
A2	0.139±0.010 a	0.484±0.149 a	103.7±11.4 ab	83.5±5.7 ab
L1	0.118±0.009 ab	0.282±0.080 bc	81.0±31.0 bc	62.5±10.8 c
M1	0.112±0.056 abc	0.219±0.053 c	90.3±17.0 abc	64.8±16.7 bc
M2	0.134±0.039 a	0.374±0.215 abc	75.0±4.0 c	66.5±21.2 bc
M3	0.055±0.016 d	0.260±0.127 c	87.6±11.4 abc	64.8±16.7 bc
M4	0.067±0.019 cd	0.204±0.046 c	106.6±3.5 ab	56±5.2 c
M5	ND*	0.332±0.048 abc	ND	56±5.2 c

The mean values of 4 replicates (agricultural) and 3 replicates (native) per accession are shown, followed by the standard deviation of the mean. The harsh conditions of the native soil prevented us to have 4 replicates for all the accessions; therefore, it was decided to normalize the number of replicates to 3 in native soil samples. Statistical analysis of root dry weight and days to flowering were done between bean accessions per soil type. ANOVA and LSD (P<0.05) tests were applied after checking for assumptions of normality and homoscedasticity. Accessions with the same letter are not significantly different.
*Accession M5 did not grow on the native soil.

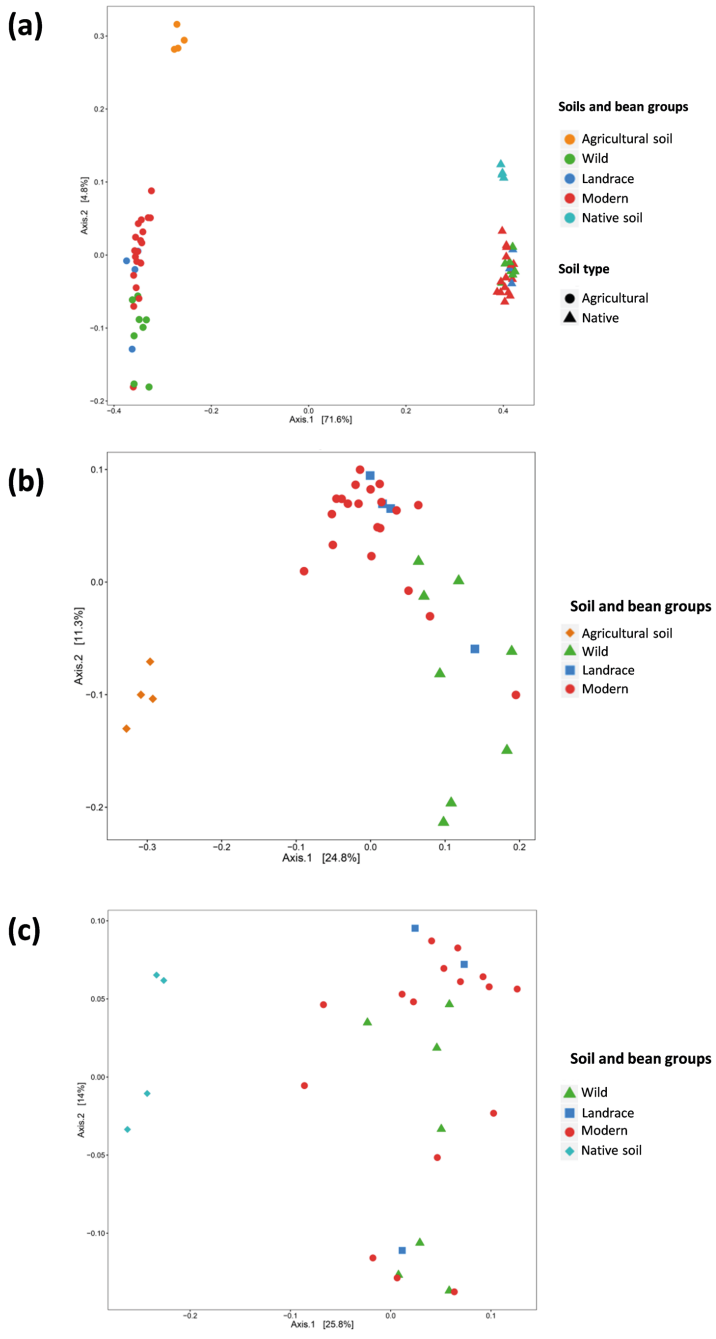


Fig. 2. Rhizosphere bacterial community structure in agricultural and native soils. Principal Coordinate Analysis (PCoA) of 16S rRNA diversity in the rhizosphere of the eight common bean accessions used in this study. (a) Rhizosphere bacterial community of common bean grown in agricultural (circles) and native (triangles) soils. Colors represent the stage of domestication and bacterial communities

from agricultural and native bulk soils. (b) PCoA including only rhizosphere bacterial communities of common bean plants grown in agricultural rhizosphere and bulk soil samples. Colors and shapes represent the stage of domestication and bacterial communities from agricultural and native bulk soils. (c) PCoA including only rhizosphere bacterial communities of bean plants grown in native rhizosphere and bulk soil samples. Colors and shapes represent the stage of domestication and bacterial communities from agricultural and native bulk soils. CSS transformed reads were used to calculate Bray-Curtis distances in (a), (b) and (c).

Profound impact of soil microbiome in the rhizosphere of common bean grown in native and agricultural soils

The observed differences in α and β diversity between the native and agricultural soils and between the eight bean accessions led us to explore more in depth the differences in taxonomic identity and relative abundance of the bacterial taxa for each soil. The most abundant bacterial phyla were Proteobacteria, Acidobacteria and Bacteroidetes in both soils. In the native soil, however, the phylum Acidobacteria showed a higher relative abundance than in the agricultural soil (Figs. S4 and S5). At phylum level, Acidobacteria and Verrucomicrobia were significantly more abundant in the native bulk soil than in the agricultural bulk soil (Welch's t-test, $P<0.05$, Bonferroni-corrected) (Fig. S6a). At class level, Acidobacterial subgroups 1, 2 and 3 were enriched in the native soil, while Acidobacteria subgroup 4 and Betaproteobacteria were more abundant in the agricultural soil (Welch's t-test, $P<0.05$, Bonferroni-corrected) (Fig. S6b). Proteobacteria and Bacteroidetes were consistently more abundant in the rhizosphere of common bean, regardless of the soil type, while Acidobacteria and Verrucomicrobia showed a consistent decrease in the rhizosphere compared to their abundance in bulk soil (Figs. 3a and b). Actinobacteria was significantly more abundant in the rhizosphere of common bean grown in the agricultural soil as compared to the bulk soil (Fig. 3a), whereas this rhizosphere effect was not observed in the native soil (Fig. 3b) (Welch's t-test, $P<0.05$, Bonferroni-corrected).

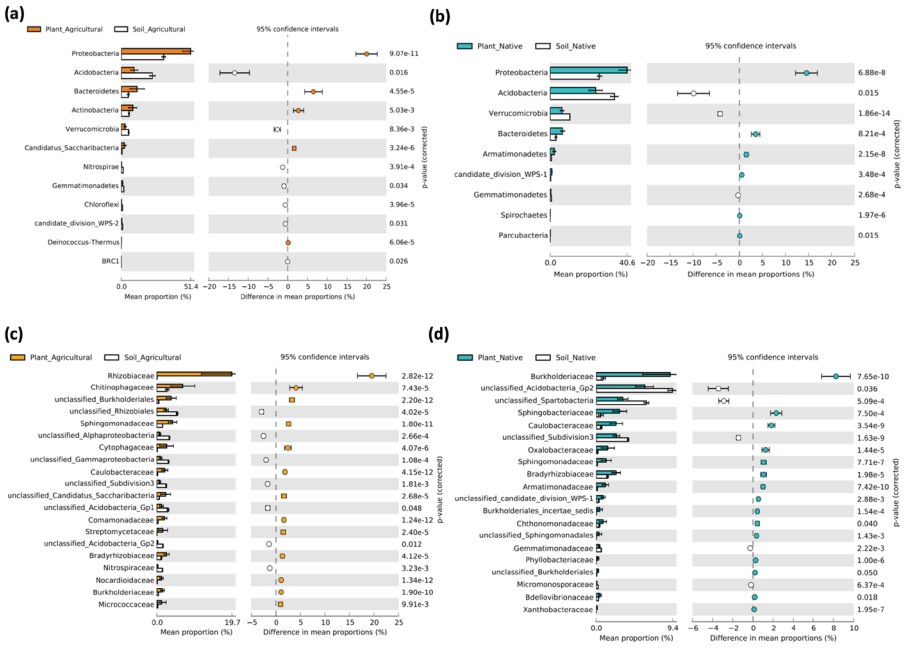


Fig. 3. Differential abundance of bacterial OTUs in agricultural and native soils. Welch's t-tests followed by Bonferroni corrections were performed between merged rhizosphere samples and merged bulk soil samples from agricultural soil and native soil at Phylum (a and c) and Class (b and d) levels. Only differentially abundant Phyla and Classes are shown.

Among the Actinobacteria enriched in the rhizosphere of all bean accessions grown in the agricultural soil, *Streptomycetaceae* and *Nocardioidaceae* were the most abundant families together with *Rhizobiaceae*, *Sphingomonadaceae, Caulobacteraceae* and *Comamonadaceae* for the Proteobacteria, and *Chitinophagaceae* and *Cytophagaceae* for the Bacteroidetes (Fig. 3c). The smaller yet significant rhizosphere effect observed for the bean accessions grown in the native soil was explained by higher relative abundances of *Burkholderiaceae*, *Caulobacteraceae*, *Oxalobacteraceae, Sphingomonadaceae* and *Bradyrhizobiaceae* for the Proteobacteria and *Sphingobacteriaceae* for the Bacteroidetes (Fig. 3d). To further dissect these differences in microbiome composition between rhizosphere and bulk soils, the abundance of the read counts was fitted to several species

abundance distribution (SAD) models. Comparison of Akaike's Information Criterion (AIC) allowed us to find the best-fit value from six distribution models. The results showed that the OTU abundance distributions in the rhizosphere of the bean accessions grown in agricultural and native soils, and the respective bulk soils, are explained by niche-based distributions (Pérez-Jaramillo *et al*. 2017) (Table S2).

Higher bacterial diversity recruited by common bean grown in agricultural soil

We performed a comparison of the bean rhizobacterial community at genus level between soil types and between wild and modern accessions in order describe how habitat expansion and agricultural soil management could have depleted or enriched rhizosphere bacterial genera. The results showed that 143 rhizobacterial genera, representing 28.7% of the total number of genera, were exclusively represented in the agricultural soil (Fig. S7a). Exclusive genera accounted for 2.3% of the total relative abundance in agricultural soil. Some of these 'exclusive' genera such as *Lysobacter* and *Aeromicrobium* accounted for 0.6% and 0.4% of the total relative abundance, respectively. Thirty-one genera including *Cytophaga* and *Acidicapsa* were exclusively found in the native soil, representing 6.2% of the total number of bacterial genera (Fig. S7a). Exclusive genera accounted for 0.2% of the total relative abundance in native soil.

Two wild accessions (A1 and A2) and two modern bean accessions (M3 and M4) were selected in order to compare the number of shared and exclusive bacterial genera in the rhizosphere. We found that in the agricultural soil, 85.9% of the rhizobacterial genera were shared between wild and modern accessions, 8.7% was exclusively found in the rhizosphere of the modern bean accessions and 5.4% was exclusively found in the rhizosphere of the wild accessions (Fig. S7b). In the native soil, a similar trend was observed, with 84.8% of the rhizobacterial genera shared between wild and modern bean accessions, 9.0% exclusively found in the rhizosphere of the two modern accessions and

6.3% in the two wild accessions (Fig. S7c). In conclusion, we found more bacterial genera in the rhizosphere of the eight bean accessions grown in the agricultural soil than in the native soil. Additionally, we found more bacterial genera in the rhizosphere of the modern bean accessions than in wild accessions irrespective of the soil type. It should be noted that the abundance of these 'exclusive' bacterial genera in the common bean rhizosphere was relatively low for both soils.

The core microbiome of common bean was represented by a small subset of microorganisms

From the total of 16,727 clustered OTUs, we found 113 OTUs consistently present in the rhizosphere of all eight bean accessions grown in the native and agricultural soils. These 113 OTUs, classified up to genus level, represented only 0.67% of the total number of OTUs but 25.9% of all the sequence reads. This core bean rhizosphere microbiome consisted of 61 Proteobacterial OTUs that made up 68.8% of the mean relative abundance with the genus *Rhizobium* as the most abundant contributor (2 OTUs, 33.4%), followed by *Bradyrhizobium* (2 OTUs, 6.7%), *Burkholderia* (3 OTUs, 4.9%), *Novosphingobium* (3 OTUs, 3.0%) and *Sphingomonas* (1 OTU, 2.2%) (Fig. 4). Other Phyla represented in the core rhizosphere microbiome were Acidobacteria (27 OTUs, 12.2% relative abundance), Actinobacteria (6 OTUs, 4.1%), Verrucomicrobia (8 OTUs, 2.5%) and Planctomycetes (5 OTUs, 1.1%). A core microbiome analysis was done also per soil type in order to dissect the specific contribution of each habitat to the overall core. For the agricultural soil, the core rhizosphere microbiome was composed of 611 OTUs representing 4.97% of the total number of OTUs and 33.07% of the reads. Proteobacteria (219 OTUs), Bacteroidetes (62 OTUs) and Actinobacteria (58 OTUs) were the three most abundant phyla within the core (Fig. S8) with again *Rhizobium* as the most abundant genus (26.7%) followed by *Dyadobacter* (3.3%) and *Streptomyces* (2.1%). In the native

soil, the core rhizosphere microbiome was composed of 812 OTUs representing 12.6 % of the total number of OTUs and 46.4% of the reads. Proteobacteria (237 OTUs), Acidobacteria (190 OTUs), Verrucomicrobia (68), Bacteroidetes (53 OTUs), Actinobacteria (48) and Chloroflexi (17 OTUs) were the most abundant phyla (Fig. S9). Within Proteobacteria, *Ralstonia* was the most abundant genus (4.6%) followed by *Burkholderia* (4.0%), *Herbaspirillum* (1.6%) and *Rhizobium* (1.2%). In the core rhizosphere microbiome of beans grown in the native soil, Acidobacteria was mainly represented by the Acidobacteria subgroups 1, 2 and 3, with 27.4% of the reads. In fact, less than 3% of the OTUs classified in the core as Acidobacteria summed up 12.7% of the total number of reads, evidencing the strong dominance of this phylum in the native soil habitat. Verrucomicrobia represented in total 6.6% of the core microbiome with most of the reads assigned as *incerta_sedis*. Finally, the genus *Mucilaginibacter* and unclassified members of the *Chitinophagaceae* family accounted for most of the abundance of Bacteroidetes representing 5.4% of the core; for Actinobacteria, small contributions mostly by unclassified Acidimicrobiales and Actinomycetales collectively accounted for a relative abundance of 3.1%. In conclusion, these comparative analyses indicated that only a small number of 113 OTUs was consistently present in the rhizosphere of all eight bean accessions grown in the agricultural and native soils, and also revealed that these OTUs represent on average more than a quarter (25.9%) of the total of 4.2 million sequence reads.

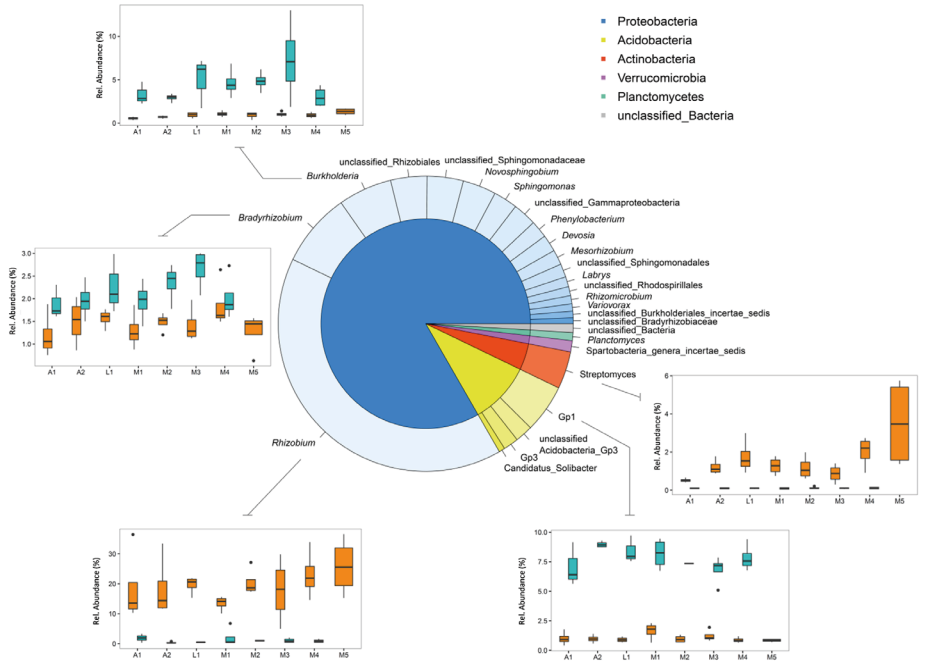


Fig. 4. Core microbiome of the rhizosphere of common bean. The different portions within the inner pie chart represent the bacterial phyla that are part of the common bean core microbiome. The outer donut plot represents the genera that are part of the core, and each genus assigned to the phylum they belong to. The size of the different pie and donut portions represents the contribution of each phylum/genus to the total relative abundance. Satellite box plots depict the relative abundance of selected genera by bean accession (A1 and A2, wild; L1, Landrace; M1 to M5, Modern) and by soil type. Cyan and dark orange colors were assigned to native soil and agricultural samples, respectively.

Higher network complexity in native soil based on co-ocurrence

Co-occurrence network analyses were performed to assess the complexity of the interactions between the rhizobacterial taxa detected in the rhizosphere of common bean grown in native and agricultural soils. The correlations between the occurrence of the rhizobacterial genera were calculated using SparCC (Friedman and Alm, 2012) followed by the graphical inference of the network and the estimation of several topological properties (Table S3). The rhizobacterial network in agricultural soil consisted of 63 nodes and 61 significant correlations, with only one negative connection between OTUs

identified as *Lysobacter* and *Ohtaekwangia* (Fig. 5a). In general, this network presented a simple structure, with four main clusters and few OTUs per cluster. For the native soil, the obtained network contained 89 nodes and 176 significant correlations, with 158 positive and 18 negative (Fig. 5b). Three main clusters were identified, with a high number of nodes per cluster and with a high number of interconnections within each cluster. Remarkably, cluster 2 was connected to the other two clusters only through negative correlations. Global networks statistics were determined for both networks (Table S3). Briefly, modularity and the number of communities were higher in the agricultural soil than in the native soil. Conversely, the average path length and the average degree were higher in the native soil. Using Betweenness Centrality (BC), we aimed to find keystone species within each network. In the agricultural soil, the highest BC values were found for the genera *Lysobacter* (OTU_136), *Rhizobium* (OTU_1), *Niastella* (OTU_10281, OTU_44 and OTU_56), *Ohtaekwangia* (OTU_69), *Terrabacter* (OTU_46), and *Arthrobacter* (OTU_886). For the native soil, the highest BC values were found for *Aquisphaera* (OTU_537, Planctomycetes phylum), two unclassified_Acidobacteria (OTU_62 and OTU_12725), an unclassified_Acetobacteraceae (OTU_175) and *Burkholderia* (OTU_45).

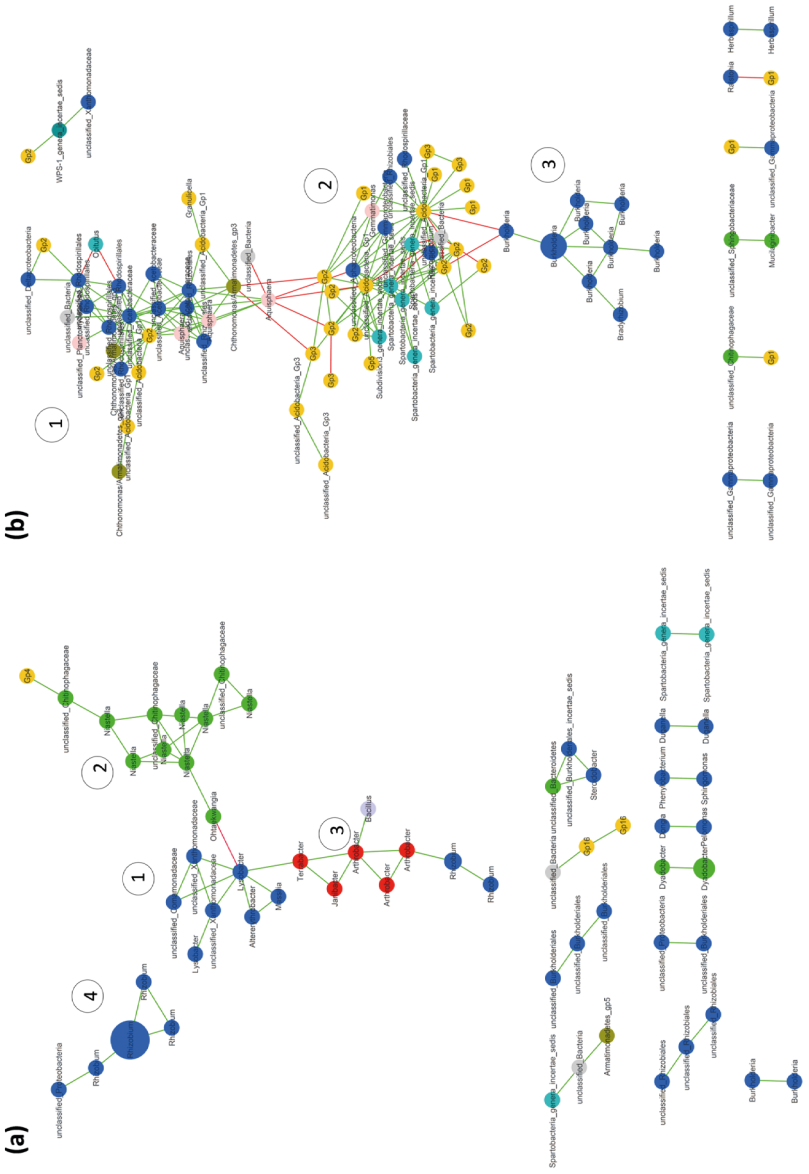


Fig. 5. Common bean rhizobacterial co-occurrence networks in agricultural and native soils. (a) Co-occurrence network of common bean rhizosphere samples in agricultural soil. Cluster 1 was composed of bacterial taxa from several classes of the Proteobacteria phylum. Cluster 2 contained exclusively bacterial taxa from the *Chitinophagaceae* family. Cluster 3 included actinobacterial taxa and one Bacillus, and cluster 4 was composed of the genus *Rhizobium*. (b) Co-occurrence network of common bean rhizosphere samples in native soil. From the three main clusters identified, two were highly abundant in nodes from the Proteobacteria phylum (1 and 3) which held negative connections to cluster 2, mainly composed of phyla Acidobacteria and Verrucomicrobia. Positive interactions are depicted as green edges and the negative interactions are depicted as red edges.

Discussion

In this study we showed that rhizobacterial diversity of both wild and domesticated common bean (*Phaseolus vulgaris*) was higher in an agricultural soil as compared to a native soil. Furthermore, species abundance analyses revealed niche-based processes for both soils suggesting selection pressures. For the agricultural soil, management practices (fertilization, addition of organic matter) are the most likely drivers of the observed differences in species abundance distributions, whereas edaphic factors, in particular low pH, are the most probable selection pressures for the native soil. Bacterial diversity is generally lower in acidic soils (Fierer and Jackson, 2006; Lauber *et al*., 2009) and pH largely determines the composition of the soil bacterial communities (Kuramae *et al*., 2012). Our results further showed that the impact of the bean genotype on rhizobacterial assembly was more prominent in the agricultural soil than in the native soil where the rhizosphere effect was much smaller and where genotype-dependent effects on rhizobacterial community composition were more homogeneous. An underlying mechanism of this minor and more homogenous rhizosphere effect is that the harsh abiotic conditions in the native soil may have affected the amount and quality of root exudates released into the soil. In the native soil used in this study, the bean plants faced a low soil pH, high aluminum concentrations and low P availability, characteristics that are common for tropical undisturbed soils (Rodrigues *et al*., 2013; Sánchez and Logan, 2013). Also, the lack of nodulation in these acidic conditions (Ferguson *et al*., 2013) could have undermined symbiotic associations for nitrogen uptake and concomitantly the growth and development of the common bean plants with an adverse effect on root exudation.

Common bean grown in the agricultural soil harbored more exclusive OTUs than bean grown in the native soil and we also found more exclusive OTUs in the rhizosphere of modern bean accessions as compared to wild accessions, irrespective of the soil type. The

genera exclusive for the agricultural soil were *Lysobacter* and *Aeromicrobium*. The genus *Lysobacter* is commonly found in agricultural soils (Puopolo *et al*., 2018) and their abundance is strongly modulated by soil type and negatively affected by low pH (Postma *et al*., 2011; Tardy *et al*., 2015). Liming is a common agricultural practice in tropical croplands to increase soil pH (Raboin *et al*., 2016) and is also typically applied in the region in Colombia where the agricultural soil used in our study was collected. Consequently, their exclusive presence in the agricultural soil might be related with the higher pH as compared to the native acidic soil. Also the exclusive genus *Aeromicrobium* prefers neutral to alkaline pH and has been previously isolated from agricultural fields (Yoon *et al*., 2005; Cui *et al*., 2007). In terms of activity, both *Lysobacter* and *Aeromicrobium* species are known to produce diverse secondary metabolites with antimicrobial properties (Miller *et al*., 1991; Hayward *et al*., 2010; de Bruijn *et al*., 2015; Gómez-Expósito *et al*., 2015) which may aid in the protection of the bean plants against biotic stress caused by soil-borne pathogens. Further experimentation is needed to decipher the impact of these 'enriched' microbes for common bean growth and health in agricultural and native habitats.

Bacteria exclusively found in the native habitat of wild relatives of crop plants, comprise representatives of *Cytophaga* and *Acidicapsa* genera. The genus *Cytophaga* is known for its cellulose-degrading capabilities and species such as *Cytophaga hutchinsonii* can be found as indigenous soil inhabitants (Wilhelm *et al*., 2017; Zhu and McBride, 2017). Their exclusive presence in the native soil may be associated with their ability to decompose complex carbohydrates such as plant litter and decaying wood, thereby contributing to carbon cycling in the undisturbed native soils. The genus *Acidicapsa*, which belongs to the phylum Acidobacteria, encompasses strictly aerobic chemo-organotrophs that are adapted to acidic conditions (Kulichevskaya *et al*., 2012; Falagán *et al*., 2017). Acidobacteria members are in general considered oligotrophs and have been

found positively associated with low soil pH (Fierer *et al*., 2007; Jones *et al*., 2009). The diversity and abundance of acidobacterial species in soil, as well as their diversity in metabolic traits, makes this phylum a potential player in soil nutrient cycling (Ward *et al*., 2009; Kielak *et al*., 2016). If these rhizobacterial genera, when re-introduced into agricultural soils, will be able to establish and survive in the rhizosphere of modern bean cultivars and, if they can, provide additional life-support functions (growth, health) for the bean plants remains to be investigated. It is important to highlight that the enriched or depleted bacterial taxa explored in this study are based on amplicon sequences that were classified up to genus level. It may be possible that bacterial species that were classified up to genus level in our analysis are absent in one of the soil types. Therefore, additional analyses that allow taxonomic resolution at the species or even strain level are needed.

A highly abundant taxonomic core was shared by all eight bean accessions in both soils. The core microbiome genera we observed included *Rhizobium*, *Bradyrhizobium, Mesorhizobium, Sphingomonas,* and *Streptomyces*. These results showed that a significant portion of the core microbiome of common bean is composed of bacterial genera with nitrogen fixing capabilities, an important feature of microbes associated with leguminous plant species. However, also for other non-leguminous plant species these rhizobacterial genera are members of the core rhizosphere microbiome (Peiffer *et al*., 2013; Yeoh *et al*., 2016; 2017). We further observed that *Rhizobium* was by far the most dominant core member in the agricultural soil while in the native soil the genera *Burkholderia, Ralstonia* and unclassified Rhizobiales were the dominant core members. These latter genera are most likely better adapted to acidic conditions in the native soil and probably responded more efficiently to root signals, such as flavonoids released by roots of common bean. *Burkholderia* species are indeed well represented in acidic soils (Stopnisek *et al*., 2014) and have been found as nodule-forming rhizobia in symbiosis

with leguminous plants (Moulin *et al*., 2001; Elliot *et al*., 2009; Lemaire *et al*., 2016) including common bean (Talbi *et al*., 2010). To form nodules, however, compatibility between *Burkholderia* spp. and the legume host is a key factor (Lemaire *et al*., 2016). In fact, common bean nodulation in tropical acid soils in South America has been associated with only a few *Rhizobium* species (Martínez-Romero *et al*., 1991; Aguilar *et al*., 2004) which were found in low abundances in the native soil. Despite the high abundance of *Ralstonia* in the native soil, a genus known to harbour soil-borne bacterial pathogens, no disease symptoms in common bean roots were evidenced in our study. *Ralstonia* species may occupy several ecological niches and have been isolated from different environments, including soil (Chen *et al*., 2001). It has been shown that tropical leguminous plants can be nodulated by *Ralstonia taiwanensis* that display functional nitrogenase activity (Chen *et al*., 2003). Nevertheless, whether the *Burkholderia* and *Ralstonia* OTUs detected here in the rhizosphere can establish symbiotic associations with common bean is not known yet.

The co-occurrence network analyses further indicated that the interactions between rhizobacterial taxa in the rhizosphere of common bean accessions grown in a native soil environment were more complex than those in an agriculturally managed soil, where the establishment of copiotrophs in the rhizosphere compartment was favored. Therefore, we hypothesize that the rhizobacterial community assembly for common bean grown in the agricultural soil is less complex and more modular than for common bean in native soil. This in turn, makes it relatively more easy for a given soil bacterial species to invade and establish in the rhizosphere of bean plants grown in the agricultural soil. Following this hypothesis, the higher rhizobacterial diversity observed for common bean in the agricultural soil may represent a less specialized microbiome. Accordingly, N-fertilization into soil has been shown to induce shifts in bacterial community composition,

promoting copiotrophs that rely on labile carbon sources (Ramirez *et al*., 2012) as well as promoting the evolution of less-mutualistic microbes (Weese *et al*., 2015). In the 'agricultural' and 'native soil' networks, we observed positive interactions between nodes, which suggest niche overlap, and we also observed negative interactions, suggesting competition or amensalism (Faust and Raes, 2012). The occurrence of phylogenetically close OTUs was in general positively correlated, forming well-differentiated clusters (Fig. 5). Accordingly, it has been shown that Acidobacteria and Verrucomicrobia phyla tend to co-occur more than expected by chance only (Barberán *et al*., 2012). In this study, we found a similar pattern in the native soil network, where cluster 2 is composed mainly of the oligotrophic phyla Acidobacteria and Verrucomicrobia. This cluster interacts negatively with clusters 1 and 3, abundant in copiotrophic bacterial genera, which presumably respond better to the common bean root exudates. Furthermore, the clustering suggests a strong niche differentiation (Faust and Raes, 2012). For instance, cluster 4 in the agricultural network is composed exclusively of rhizobial OTUs, with no interactions with other clusters. Similarly, cluster 2 consisted mainly of Bacteroidetes that may represent the rhizobacterial hub involved in degradation of complex polymers (Thomas *et al*., 2011; Berlemont and Martiny, 2015). Also cluster 3 in the native soil, mainly composed of *Burkholderia*, might represent a specific hub of nodule-forming rhizobia (Talbi *et al*., 2010). Whether these hubs represent distinct functional groups remains to be investigated by metagenomics and trait-based bioassays.

Conclusions

Our study showed that the transition of common bean from a native soil to an agricultural soil led to a gain of rhizobacterial diversity. We found a low diverse but highly abundant core microbiome which resembles that of other plant species, suggesting a homogenization of rhizobacterial diversity of plants grown in different agricultural

landscapes. It is important to note that the core microbiome analysis presented here is a descriptive analysis based solely on taxonomy, and that functional traits should be taken into account in future analyses for better insight into the impact of habitat expansion on trait-based microbiome assembly (Lemanceau *et al*., 2017). The network structure was simpler in agricultural soil as compared to native soil, which again may reflect the process of biotic homogenization. In this study, we also aimed for the identification of microbes that were lost as a consequence of domestication and habitat expansion of common bean. Indeed, several bacterial genera were exclusively found in the native soil and also as an exclusive member of the rhizosphere of wild bean accessions. These bacterial genera were low-abundant members of the rhizobacterial community. Conversely, the number of bacterial taxa exclusively found in the agricultural soil was considerably higher. The proportion of depleted bacterial genera appears to be overcompensated in the agricultural soil by the number of "gained microbes", many of which were highly abundant in the rhizosphere of all eight common bean accessions. On the other hand, this increased bacterial diversity in the agricultural soil might also correspond to a less specialized microbiome. To what extent these "enriched" and "depleted" bacterial genera have an impact on plant growth and health is not yet known and subject of future experiments. It is important to emphasize that the number of agricultural and native soils tested should be further expanded to resolve if the significant changes we observed between the two soils tested in our study can be extrapolated as general trends in rhizobacterial shifts during domestication.

Acknowledgements

J.E.P-J was financially supported by the Department of Science, Technology and Innovation of Colombia - COLCIENCIAS through the doctoral grant 568-2012-15517825. J.M.R. and V.J.C. were supported by the Dutch STW-program "Back to the Roots" and R.M. by CNPq 443112/2014-2. We are grateful to H. A. Pérez and J. A. Pérez

for assistance in the selection and collection of the native soil for the greenhouse experiments.

Supplementary materials

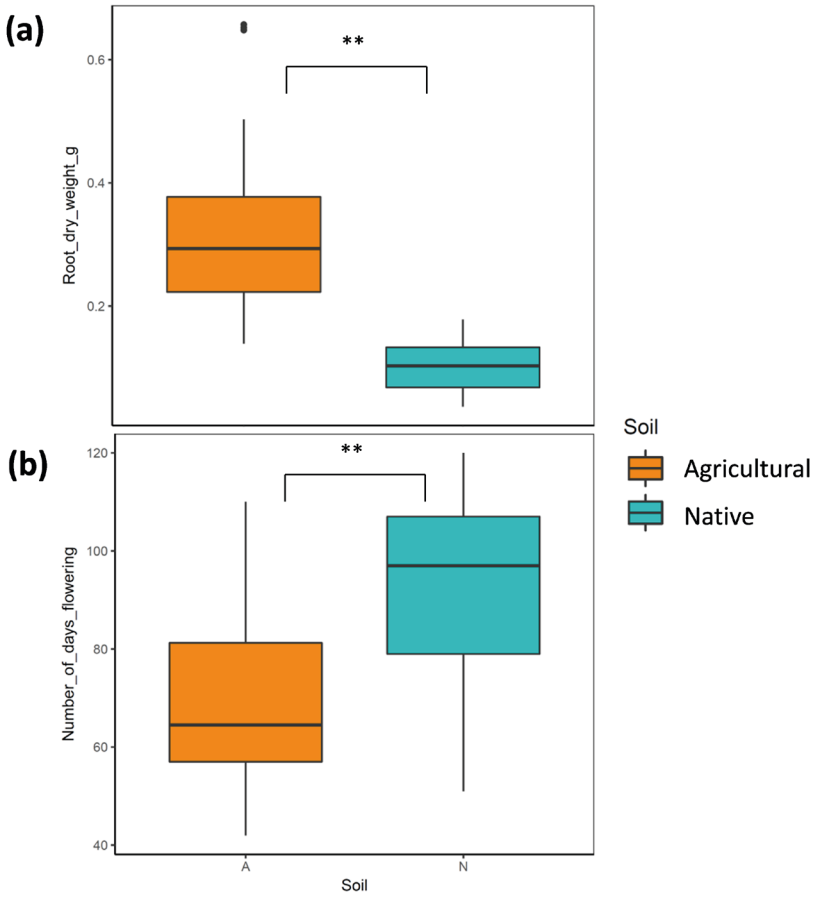


Fig. S1. Phenological traits of the common bean accessions grown in native and agricultural soils. Data for root dry weight (a) and the number of days to reach the flowering stage (b) of common bean accessions grown in agricultural and native soils was merged per soil type and used to built box plots. Subsequently, kruskal and Wallis test was applied followed by a Bonferroni correction. (** significantly different with a $P<0.05$)

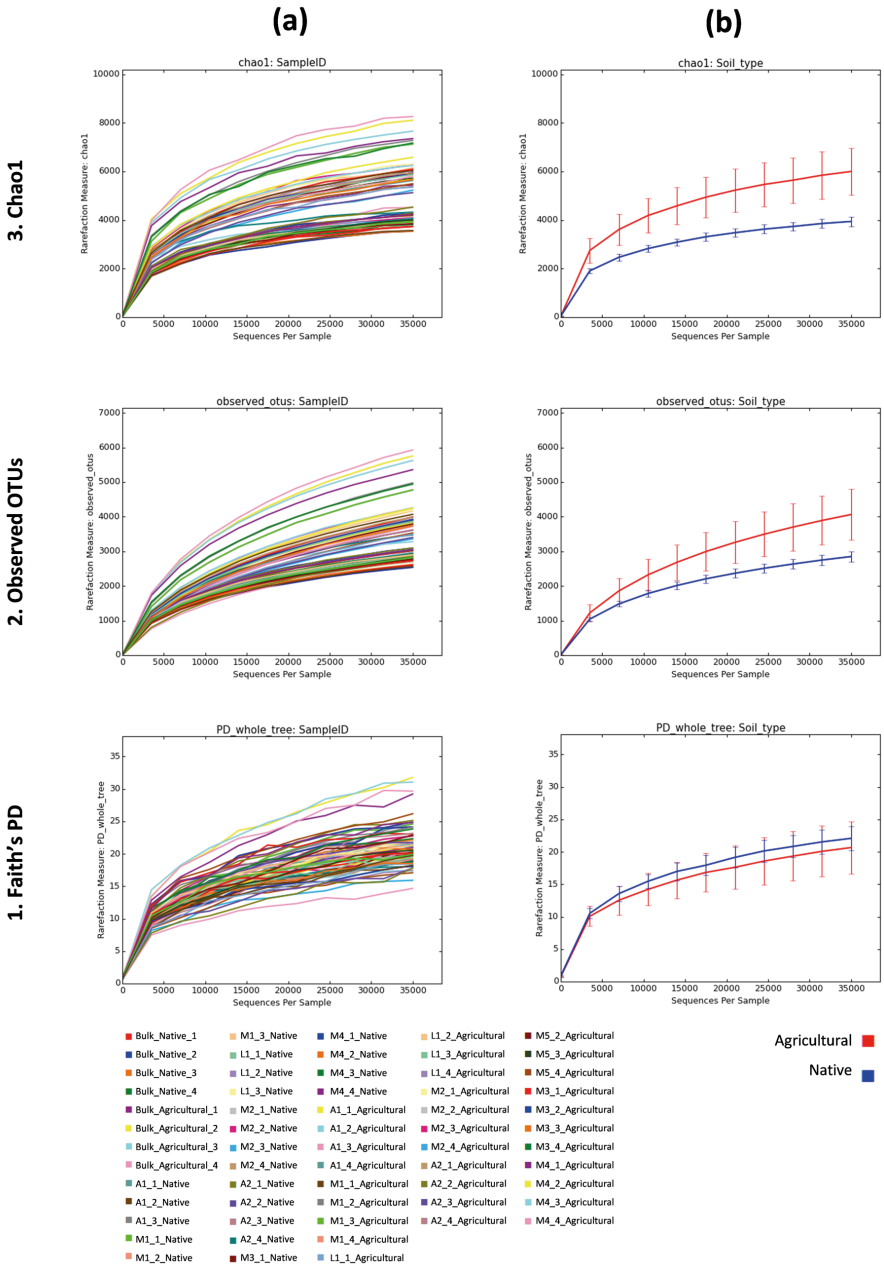


Fig. S2. Rarefactions curves for chao1, observed OTUs and phylogenetic diversity metrics. In the rows, the rarefaction curves are shown by index as follows: 1. Faith's phylogenetic diversity; 2. Number of observed OTUs; 3. chao1 index for estimated OTUs. In the rows, rarefaction curves are shown by (a) Sample; (b) Soil type.

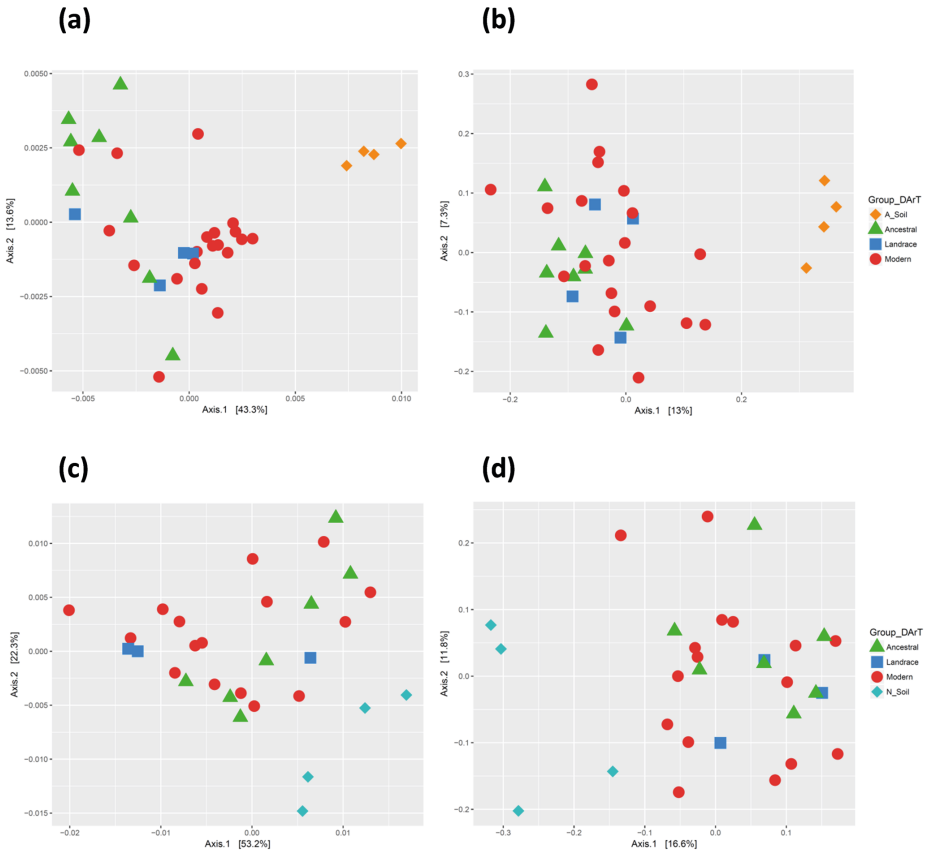


Fig. S3. Rhizosphere bacterial community structure using weighted Unifrac metrics in agricultural and native soils. Principal Coordinate Analysis (PCoA) of 16S rRNA diversity in the rhizosphere of the eight common bean accessions used in this study. (a) weighted and (b) unweighted unifrac metrics of agricultural rhizosphere and soil samples and (c) weighted and (d) unweighted unifrac metrics of native rhizosphere and soil samples.

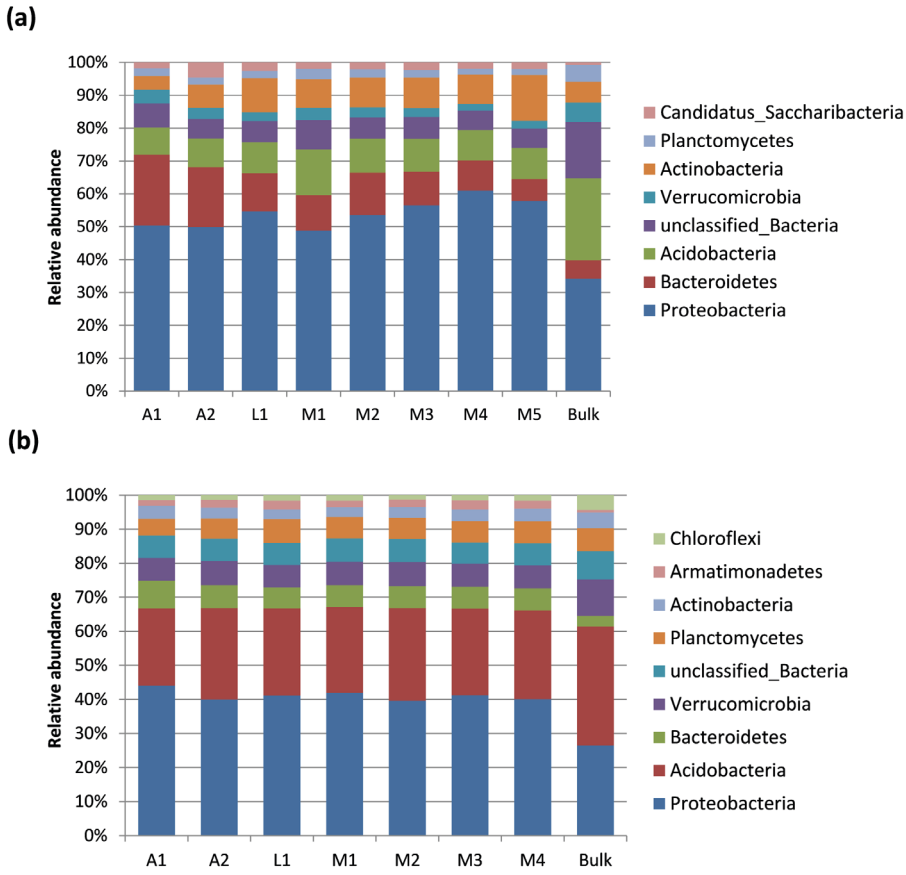


Fig. S4. Relative abundance of the most abundant bacterial phyla in agricultural and native soils. Bar graphs of the relative abundance of the most abundant bacterial phyla in the agricultural (a) and in the native (b) soil are shown. Only phyla with a total relative abundance higher than 1% are depicted in the graphs.

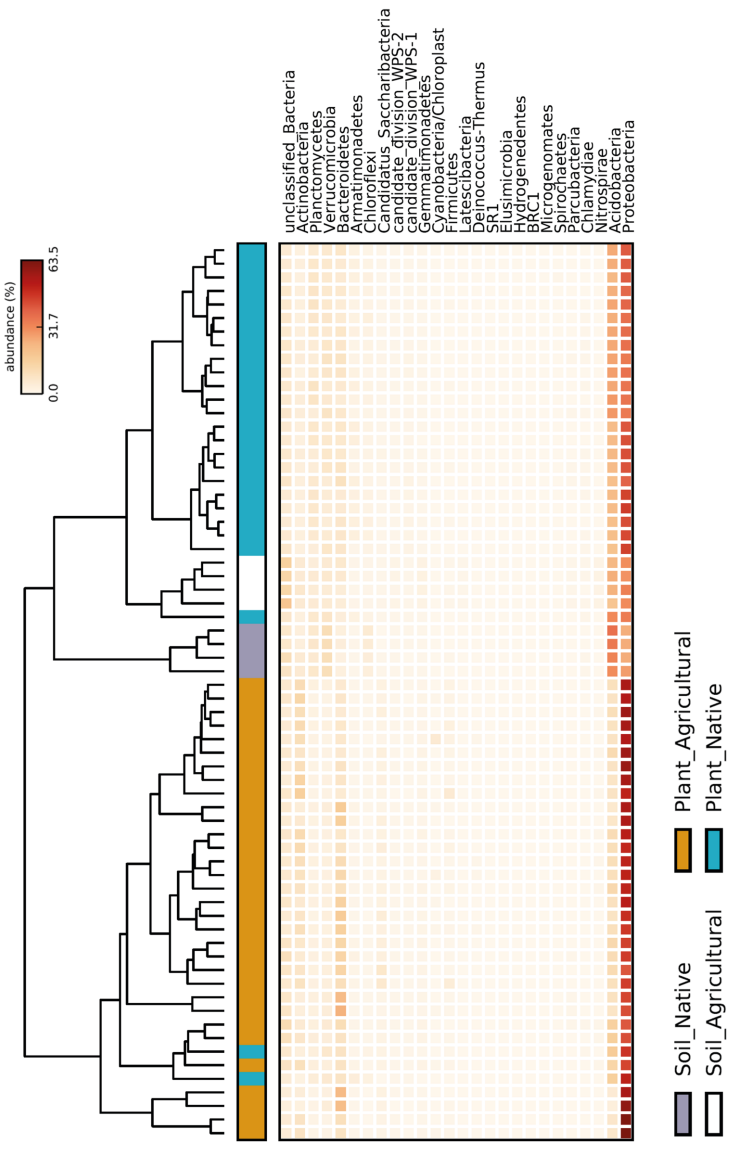


Fig. S5. Heat map of the relative abundance of all the bacterial phyla in agricultural and native soils. The columns indicate the common bean accessions clustered through a dendogram built with an average neighbor method (UPGMA) and the rows shows the bacterial phyla observed in rhizosphere and bulk soil samples. The darker the red squares in the heat map the more abundant the phylum of the specific row. Color codes for the samples are indicated in the lower section of the graph.

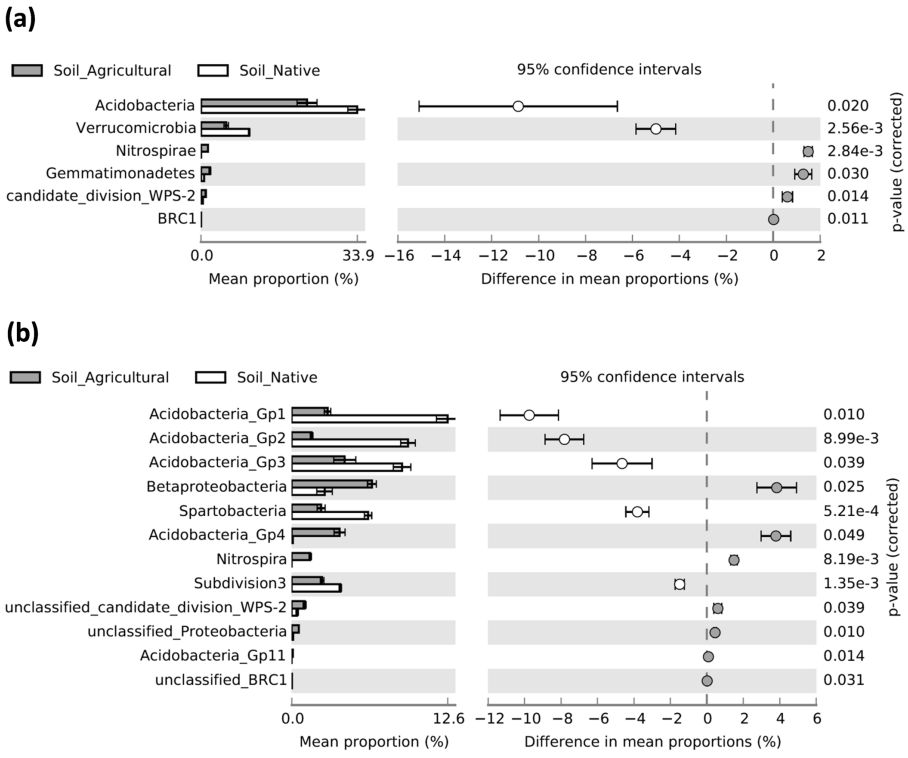


Fig. S6. Differential abundance of bacterial phyla and classes between bulk soil samples of the agricultural soil and the native soil. Welch's t-tests followed by Bonferroni (P<0.05) corrections were performed at Phylum (a) and Class level (b). Only Phyla and Classes differentially abundant are shown.

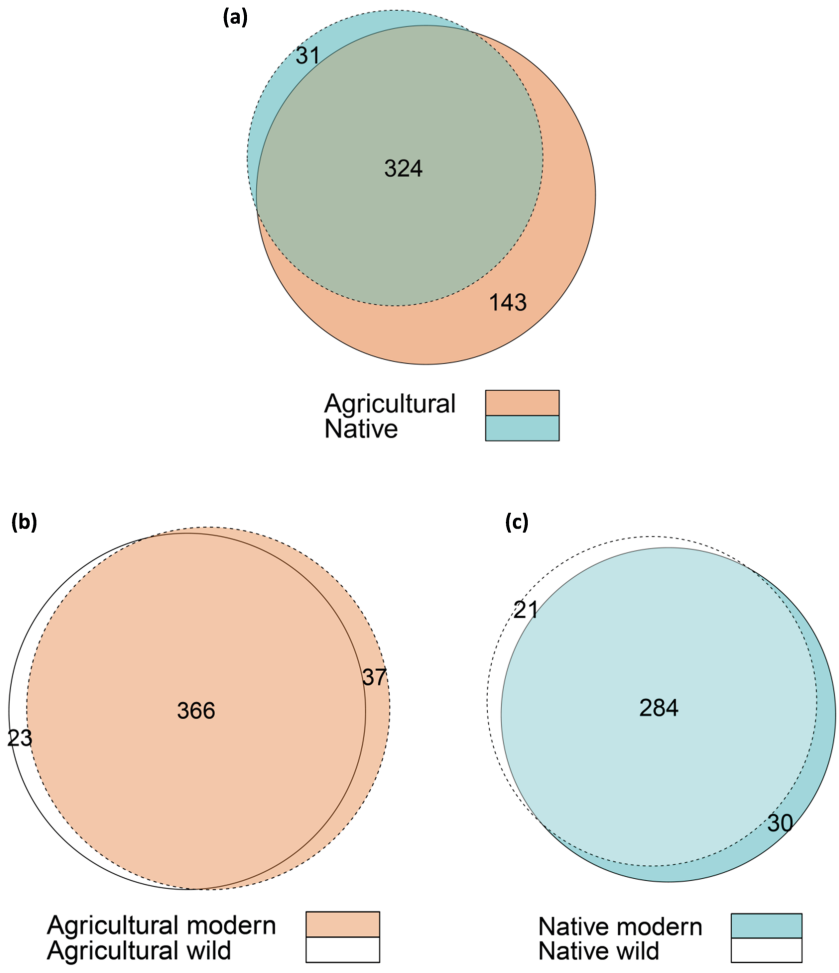


Fig. S7. The enriched and depleted bacteria in the rhizosphere of wild and modern common bean grown in native and agricultural soils. Area-proportional Euler diagrams were built to depict the exclusive and the shared genera. (a) Number of genera shared between agricultural and native soils is depicted within the intersection while the number of genera exclusive to each soil type can be seen out of the intersection zone. The genera exclusive to the native soil are visible in the cyan colored area, while genera exclusive to the agricultural soil are visible in the dark orange colored area. (b) Euler diagram between modern accessions and wild accessions in the agricultural soil. The genera exclusive to the modern accessions is visible in the dark orange colored area, while genera exclusive to the wild accessions is visible in the white area. (c) Euler diagram between modern accessions and wild accessions in the native soil. The genera exclusive to the modern accessions are visible in the cyan colored area, while genera exclusive to the wild accessions are visible in the white area.

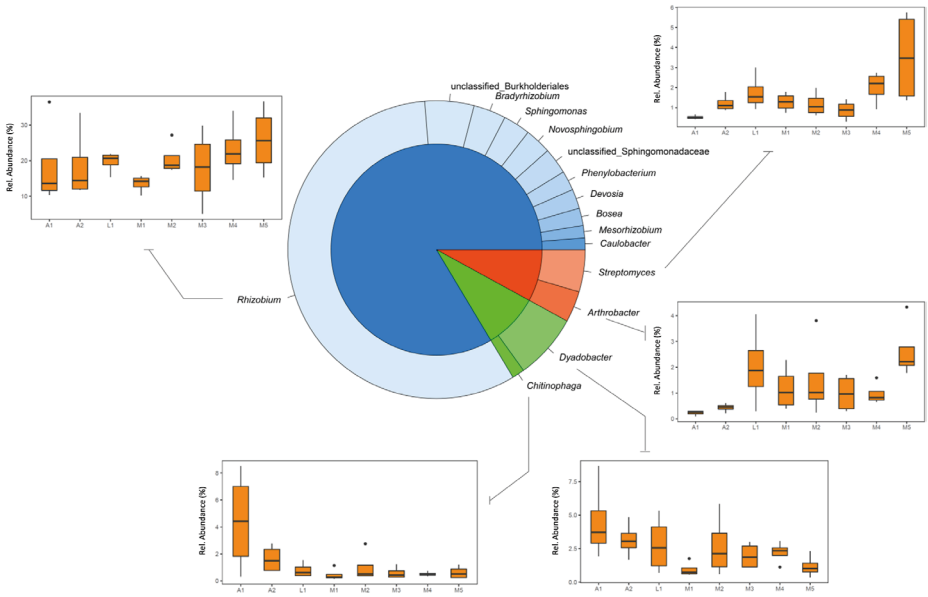


Fig. S8. Core microbiome in the rhizosphere of common bean grown in agricultural soil. The different portions within the inner pie chart represent the bacterial phyla. The outer donut plot represents the genera that are part of the core. The size of the different pie and donut portions represents the contribution of each phylum/genus to the total relative abundance. Satellite box plots depict the relative abundance of selected genera by accession in the agricultural soil.

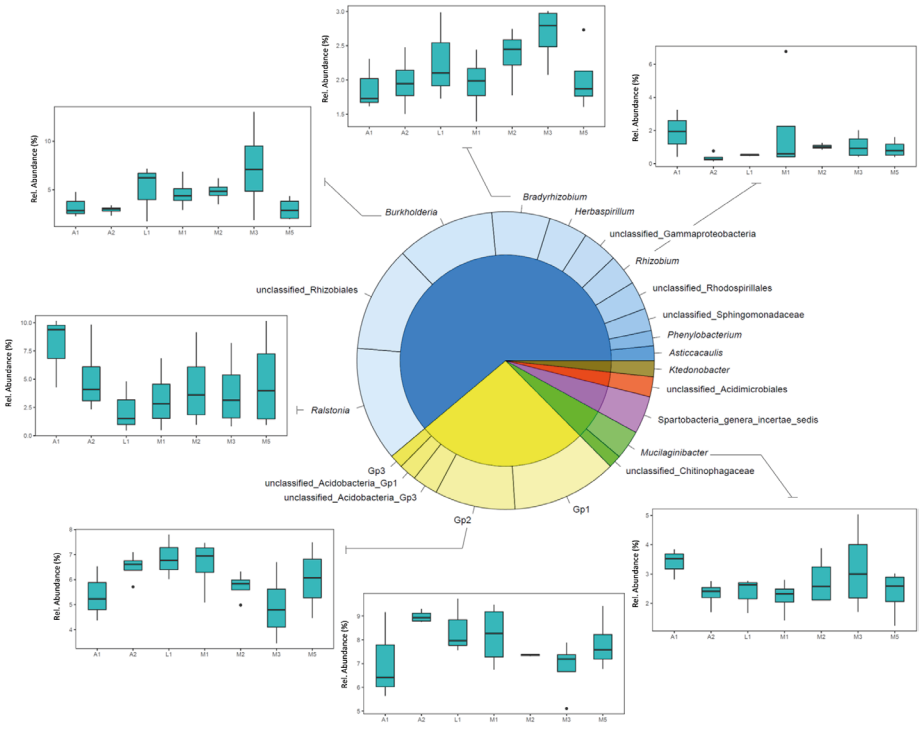


Fig. S9. Core microbiome in the rhizosphere of common bean grown in native soil. The different portions within the inner pie chart represent the bacterial phyla. The outer donut plot represents the genera that are part of the core. The size of the different pie and donut portions represents the contribution of each phylum/genus to the total relative abundance. Satellite box plots depict the relative abundance of selected genera by accession in the native soil.

Table S1. Sequence characteristics obtained in the agricultural soil. In the table the raw read and after processing read counts are presented as well as quality characteristics of the raw reads.

Soil type	Sample	Total Bases	Raw read count	GC (%)	Q20 (%)	Q30 (%)	Processed read count
Agricultural	A1_1	63363676	211312	54.71	86.63	74.13	56996
Agricultural	A1_2	78985945	263296	55.18	87.15	75.19	77864
Agricultural	A1_3	69257541	231142	54.3	87.72	76.1	70601
Agricultural	A1_4	72760654	242636	54.82	87.79	76.02	69084
Agricultural	A2_1	69773872	232626	55.02	86.96	75.01	67646
Agricultural	A2_2	58210761	194134	54.35	88.36	76.93	60724
Agricultural	A2_3	76327328	254550	54.79	87.61	75.93	75603
Agricultural	A2_4	70190462	234008	55.13	86.55	74.44	67151
Agricultural	L1_1	72057959	240260	54.87	87.57	75.82	72136
Agricultural	L1_2	64508620	215002	55.81	87.01	74.95	63095
Agricultural	L1_3	68590650	228646	55.27	87.21	75.28	66964
Agricultural	L1_4	71258796	237562	55.41	87.34	75.6	72020
Agricultural	M1_1	71348421	237868	55.22	88	76.53	71859
Agricultural	M1_2	65375690	217894	55.94	86.24	73.77	58752
Agricultural	M1_3	67841335	226132	55.83	87.02	74.98	63630
Agricultural	M1_4	65532781	218424	55.57	86.86	74.81	63673
Agricultural	M2_1	67221432	224064	55.6	87.34	75.58	67676
Agricultural	M2_2	62678239	209020	55.15	87.26	75.32	61303
Agricultural	M2_3	66567882	221854	55.88	87.19	75.32	67063
Agricultural	M2_4	60348755	201286	54.78	87.97	76.36	61242
Agricultural	M3_1	59335191	197832	55.59	87.46	75.89	60023
Agricultural	M3_2	68861207	229586	55.1	88.35	76.94	70813
Agricultural	M3_3	68907707	229716	55.49	87.93	76.19	71553
Agricultural	M3_4	68367576	227898	55.67	87.18	75.22	68092
Agricultural	M4_1	74886328	249604	55.37	87.99	76.57	79568
Agricultural	M4_2	74054373	246918	55.4	87.51	75.82	73404
Agricultural	M4_3	74769091	249182	55.69	87.19	75.23	79061
Agricultural	M4_4	66904055	223064	55.4	87.95	76.29	69317
Agricultural	M5_1	62075633	206888	55.89	86.95	74.93	60638
Agricultural	M5_2	72482267	241660	55.61	87.82	76.35	74295
Agricultural	M5_3	61109216	203688	55.7	87.48	75.84	62822
Agricultural	M5_4	71365327	237914	55.51	87.73	76.17	72160
Agricultural	Bulk_soil_1	61607085	205346	56.88	86.33	73.98	59869
Agricultural	Bulk_soil_2	65310513	217716	56.66	85.85	73.09	58020
Agricultural	Bulk_soil_3	55162959	183890	56.6	86.69	74.51	51738
Agricultural	Bulk_soil_4	60880405	202966	56.4	86.27	73.96	55876

Table S1 (Continuation). Sequence characteristics of samples obtained in the native soil. In the table the raw read and after processing read counts are presented as well as quality characteristics of the raw reads.

Soil type	Sample	Total Bases	Raw read count	GC (%)	Q20 (%)	Q30 (%)	Processed read count
Native	A1_1	69790674	232642	55.48	87.61	75.83	68643
Native	A1_2	59019732	196740	55.3	87.94	76.32	59343
Native	A1_3	64931567	216392	56.16	86.45	73.82	60661
Native	A2_1	57934044	193128	56.32	86.58	74.47	55570
Native	A2_2	61055744	203522	56.51	87.04	74.95	60234
Native	A2_3	67056980	223502	55.93	86.85	74.57	63884
Native	A2_4	58518207	195056	55.75	87.08	75.01	56197
Native	L1_1	65521365	218470	56	87.59	75.83	65232
Native	L1_2	64785862	215952	55.81	87.13	75.17	65472
Native	L1_3	71386826	238006	56.23	87.42	75.58	70322
Native	M1_1	63411035	211380	55.54	87.88	76.22	62462
Native	M1_2	62074225	206932	56.22	87.78	76.08	63354
Native	M1_3	71515302	238420	56.3	86.88	74.78	68664
Native	M1_4	54123563	180468	55.92	87.98	76.29	55091
Native	M2_1	67140790	223824	56.07	87.71	76.01	67260
Native	M2_2	66608745	222070	56.26	87.66	75.92	65336
Native	M2_3	58726189	195734	56.03	86.89	74.83	56402
Native	M2_4	70057132	233546	55.9	87.05	75.05	69406
Native	M3_1	63700948	212352	55.9	87.25	75.32	63248
Native	M3_2	58356921	194520	56.26	86.7	74.5	55418
Native	M3_3	64816913	216052	56.18	86.53	74.28	60174
Native	M3_4	52757955	175792	55.73	85.03	71.51	45340
Native	M4_1	61945280	206504	55.92	87.14	75.06	60709
Native	M4_2	66279461	220998	55.67	87.04	74.99	63633
Native	M4_3	62281114	207616	55.92	87.18	75.18	59404
Native	M4_4	52537073	175068	56.84	82.13	67.37	35547
Native	Bulk_soil_1	51904334	173074	57.05	86.92	74.74	50134
Native	Bulk_soil_2	57678950	192304	57.01	86.5	74.18	54399
Native	Bulk_soil_3	64011567	213402	56.82	85.94	73.51	59298
Native	Bulk_soil_4	61453027	204890	56.77	86.45	74.16	57379

Table S2. AIC values for six rank abundance distribution models to test niche neutral models in native soil samples. The lowest AIC (Akaike Information Criterion) value for each sample represents the best fit model which is shown in bold.

	AIC					
Sample	**Broken-stick**	**Pre-Emption**	**Log-Normal**	**Zipf**	**Zipf-Mandelbrot**	**ZSM**
A1_1	109193.1	82138.6	**16519.6**	23002.6	22313	27723.2
A1_2	94184.6	69323.1	**12489.5**	17758.8	15560.7	25791.4
A1_3	75917.1	53531.6	**12942.3**	22601.2	nd	31087.8
A2_1	63993.3	46570.4	**13574.7**	22370.6	nd	29473.8
A2_2	71558.9	51409.5	**13565.3**	22980.2	nd	29706
A2_3	97215	73862	**16099**	22244	21654	27916.2
A2_4	67463.6	47477.7	**12704.8**	21889.1	nd	28406.8
L1_1	84656.2	57675.6	**13392.4**	24235.9	nd	29305.8
L1_2	76243.7	53640.4	**13452.2**	24314.4	nd	31883.8
L1_3	83374	53282	**14091**	28556	nd	28726.4
M1_1	102404	78828.6	**15112.9**	18962.9	17328.2	30968.2
M1_2	76043.8	51113.2	**12713**	23954.3	nd	30633.6
M1_3	86084.7	58044.4	**14829**	27170	nd	31285
M1_4	63313.4	45584.3	**12505.8**	21131.8	nd	28744.8
M2_1	81233.1	53271.9	**14479.4**	27657.3	nd	28957.2
M2_2	70881	47165.9	**14569**	27979.2	nd	31003.2
M2_3	82526.2	59967.7	**12385.5**	18537.5	nd	26264.8
M2_4	95319.6	64444	**12840.3**	23598.4	nd	27933.6
M3_1	83007.8	55441.1	**13188.9**	24134.3	nd	29499.2
M3_2	80158.3	60099.8	**13045.7**	18270.1	16701.8	29880.6
M3_3	82194.7	59553.1	**13634.7**	22082.8	nd	33823
M3_4	69381	50489	**10947.5**	15735.1	12638.3	25404.7
M4_1	65597.2	45892.6	**13894.5**	25206.3	nd	33552.8
M4_2	98371.9	75718.3	**16064.2**	21553.6	21041.2	29379
M4_3	75332.4	54770.3	**13519.5**	22214.1	17947.7	28705.4
M4_4	38329.1	29292.2	**11450.2**	17028.1	nd	27612.2
Bulk_soil_1	61291.3	41596.6	**12405.9**	21171.8	nd	24761.1
Bulk_soil_2	70455.2	45471.4	**12031.8**	21858.8	nd	24218.5
Bulk_soil_3	72354.2	47488.2	**12155.4**	22969	nd	25293
Bulk_soil_4	70091.6	45407.7	**12925**	24283.3	nd	25246.2

Table S3. Co-occurrence network properties of 16S rDNA rhizobacterial reads of common bean rhizosphere in agricultural and native soil

Network properties	Common bean Agricultural soil	Common bean Native soil
Number of nodes[a]	63	89
Number of edges[b]	61	176
Positive edges[c]	60	158
Negative edges[d]	1	18
Modularity	0.822	0.584
Number of communities	17	13
Network diameter	11	11
Average path length	3.87	4.06
Average degree	0.968	1.978
Average clustering coefficient	0.5	0.551

[a]Bacterial taxa (genus level) with at least one significant (P<0.01) and strong (SparCC > 0.8 or < -0.8) correlation.

[b]Number of connections/correlations obtained by SparCC analysis;

[c]SparCC positive correlation (> 0.8 with $P < 0.01$);

[d]SparCC negative correlation (< -0.8 with $P < 0.01$)

Chapter 5

The wild side of plant microbiomes

Juan E Pérez-Jaramillo, Victor J Carrión, Mattias de Hollander, Jos M Raaijmakers

Microbiome 6:143 (2018)

https://doi.org/10.1186/s40168-018-0519-z

Abstract

Plants rely on their microbiome for a number of life-support functions including nutrient acquisition and protection against (a)biotic stress factors. For crop plants, however, the process of domestication may have adversely impacted the composition and functions of the associated microbiota, thereby undermining their beneficial effects on plant growth and health. Here, we conducted a meta-analysis to resolve if and how plant domestication affected the composition of the root-associated microbiome. For different plant species, we observed significant and consistent differences in the abundance of Bacteroidetes, Actinobacteria and Proteobacteria. Potential causes and consequences of these microbiome shifts following plant domestication are discussed.

Keywords: rhizosphere and root microbiome, wild relatives, modern cultivars, Bacteroidetes

Introduction

In the search for new strategies to engineer "healthy microbiomes" of plants and humans, considerable attention is given to coevolutionary signatures of host-microbe interactions and mechanisms involved in microbiome assembly and activity (Mueller and Sachs, 2015; Schnorr *et al*., 2016; Crittenden and Schnorr, 2017). For example, comparative analyses of the human microbiome revealed a higher abundance of Bacteroidetes in the gut of hunter-gatherer populations of rural communities in non-industrialized regions than in the gut of Westernized populations, a distinct divergence that appears to be associated with differences in the content of starch, fiber and plant polysaccharides in the food (Schnorr *et al*, 2014; Gomez *et al*., 2016). Similarly, shifts in gut microbiome composition in captive mammals as compared to their wild counterparts have been associated with a loss of dietary fiber and a potential increase in protein consumption (Clayton *et al*., 2016; McKenzie *et al*, 2017). Interestingly, one of the most relevant changes in the gut microbiome of mammals in captivity is an increase in the relative abundance of the genus *Bacteroides* and a decrease of the genus *Prevotella*, both from the Bacteroidetes phylum, a pattern that has also been observed in Westernized humans (McKenzie *et al*, 2017). For plants, several studies have suggested that domestication altered the composition of the root microbiome with an adverse effect on the association with symbiotic nitrogen-fixing rhizobia and mycorrhizal fungi (Pérez-Jaramillo *et al*., 2016). For instance, Kiers *et al*., (2007) showed that older soybean cultivars had a higher yield difference ratio, i.e. the ability of soybean cultivars to reach their full symbiotic potential in the presence of a mix of rhizobial strains with different symbiotic effectiveness, as compared to newer soybean cultivars. Similarly, it has been shown that wild ancestors and primitive landraces of wheat, breadfruit and maize can benefit more from mycorrhizal symbiosis than modern cultivars (Hetrick *et al*., 1992, 1995; Sangabriel-Conde *et al*., 2012; Xing *et al*., 2012). To date, however, the impact of plant domestication on the vast majority of other root-

associated microorganisms is not well understood. In a recent study, we revealed that the rhizosphere microbiome of wild relatives of common bean (*Phaseolus vulgaris*) harbored a higher abundance of Bacteroidetes, while the root microbiome of modern bean accessions was dominated by bacterial families belonging to the Actinobacteria and Proteobacteria (Pérez-Jaramillo *et al*., 20170). Also studies on other plants species, including Arabidopsis (Schlaeppi *et al*., 2014), sugar beet (Zachow *et al*., 2014), barley (Bulgarelli *et al*., 2015) and lettuce (Cardinale *et al*., 2015), suggested that domestication led to compositional changes in the root microbiome. To investigate if these effects of domestication cause similar shifts in microbiome composition for multiple plant species, we set out a meta-analysis of the root microbiome of various crop plants and their wild relatives. The specific objectives of this computational 'walk on the wild side' were to: i) determine differences and patterns in root microbiome composition between wild relatives and their domesticated counterparts, and ii) identify the relative abundance of specific taxa within the Bacteroidetes phylum for crop plants and their wild relatives.

Methods

Processing of the sequences

Per sample fastq files of 16S metagenome amplicon sequences were kindly provided by the authors of the different studies (Schlaeppi *et al*., 2014; Zachow *et al*., 2014; Bulgarelli *et al*., 2015; Cardinale *et al*., 2015; Leff *et al*., 2017; Pérez-Jaramillo *et al*., 2017). The reads were quality filtered for single end reads with sickle (Joshi *et al*., 2011), and bases below phred score 36 and shorter than 100bp were trimmed. Only high quality filtered reads were mapped to full length 16S sequences from the Silva 119 release (Quast *et al*., 2013) using the usearch global algorithm implemented in VSEARCH version 1.9.6 (Rognes *et al*., 2016). The alignment results were directly converted to BIOM format using biom version 2.1.5 (McDonald *et al*., 2012). Consensus/majority taxonomy was

added as metadata to the biom file. Finally, all BIOM files of each dataset were merged using Qiime version 1.9.1 (Caporaso *et al*. 2010). The Silva 119 reference phylogenetic tree provided by Qiime (clustered at 97%) was filtered using the Qiime command *filter_tree.py* to keep Bacteroidetes taxa which were present only in wild plants. Subsequently, we built phylogenetic trees using the Phyloseq package in R, and for graphic purposes only branches with a relative abundance higher than 0.1% from the total amount of reads were kept.

Description of plant and soils used in the studies

In the study by Pérez-Jaramillo *et al*., (2017), the wild and modern common bean accessions (*Phaseolus vulgaris*) were cultivated in the same soil in a pot trial under the same climatic conditions followed by characterization of the rhizosphere microbiome composition. For Cardinale *et al*. (2015), it is described in the main text that the experiments with *Lactuca serriola*, wild relative of lettuce, and four subspecies of *L. sativa* were done in the field in an experimental farm in Austria, followed by characterization of the rhizosphere microbiome composition. Nevertheless, the description of soil characteristics is not provided and therefore it is not possible to describe growth conditions for the plants neither. In the study by Zachow *et al*. (2014), wild beet plants were collected in the drift line at the Mediterranean Sea coast in Slovenia. From the same region, soil from the coastal drift line was collected and used under unspecified greenhouse conditions in order to grow domesticated beet in the same soil than wild beet. For the study by Schlaeppi *et al*. (2014), several field experiments and greenhouse experiments were done and four different types of soil were used. The root microbiome composition was characterized for *Arabidopsis thaliana*, *Cardamine hirsuta, A. halleri* and *A. lyrata*, while the rhizosphere microbiome composition was characterized for *A. thaliana* and *C. hirsuta*. The latter is an *Arabidopsis* relative species which diverged

~35 Mya and is phylogenetically the most distant species. Finally, in the study by Bulgarelli *et al*. (2015), the microbiome composition of root and rhizosphere compartments of wild barley (*Hordeum vulgare* spp. *spontaneum*), a landrace and a modern variety of barley (*H. vulgare* spp. *vulgare*) were characterized. For this, two pot trials were performed with the three plant accessions in soils that were collected in the same location in two different years. All the information about plant accessions, soil type and experimental conditions are described in Tables S1 and S2.

Statistical analysis

In order to compare the different datasets we rarefied the OTU table up to 500 reads, which was the sequencing depth that allows us to work with most of the data sets available. All the data sets were included except for the data of Leff *et al*. (2017), for which sequencing depth, after processing with the method described above, did not reach the threshold implemented. For Alpha diversity metrics, the command *alpha_diversity.py* in Qiime was applied and the output files were retrieved and plotted in R using the package ggplot2 (v.2.0.0) (Wickham, 2009). As we did not observe significant differences in alpha diversity indexes between wild and domesticated accessions of the same plant species, the data was merged per plant species in order to illustrate exclusively differences between compartments (root/rhizosphere). For beta-diversity calculations, a Bray–Curtis dissimilarity matrix was calculated and used it to build Principal Coordinate Analyses and Permutational multivariate analyses of variance (Adonis function) were performed to evaluate the significance of the variables tested, both retrieved from Phyloseq (v.1.10) (McMurdie and Holmes, 2013) and Vegan (v.2.4-4) (Oksanen *et al*., 2017). For the OTU level analysis, the function calculateEffectiveSamples from the metagenomeSeq R package (v.1.12) (Paulson *et al*., 2017) was applied to the filtered OTU table and features with less than the average number of effective samples in all features

were removed. For the analysis at OTU level, we used normalized tables applying a cumulative-sum scaling normalization. Then, a Zero-Inflated Gaussian Distribution Mixture Model was applied using the fitZig function from metagenomeSeq. With the coefficients from the model, we applied moderated t-tests between accessions using the makeContrasts and eBayes commands retrieved from the R package Limma (v.3.22.7) (Ritchie *et al*., 2015). Obtained P-values were adjusted using the Benjamini–Hochberg correction method. Differences in the abundance of taxa between accessions were considered significant when adjusted P-values were lower than 0.05 at OTU level. Treemap (v.3.7.3) was used to visualize the significantly abundant OTU's, the taxonomy, the adjusted *P*-value and per mil relative abundance in bubble graphs, in which the size of the bubbles indicates the relative abundance per hundred of the raw read counts.

Results and discussion

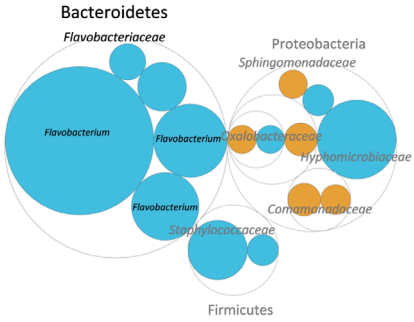
In this study, we retrieved the raw 16S rDNA sequences from 6 independent common garden experiments with a total of 9 plant species and adopted the same computational pipeline to assess the root/rhizosphere bacterial community composition (Table S1, Table S2). Regarding the analysis of the Arabidopsis root microbiome by Schlaeppi *et al.* (2014), our comparison was made based on divergence time estimates with *Cardamine hirsuta* considered as the 'ancient/wild' species and members of the genus *Arabidopsis* as the 'modern/evolved' counterpart.

First, we observed marked differences in the diversity of bacterial communities associated with roots of the different plant species, which were largely explained by the study (29.1%, PERMANOVA, $P<0.001$) (Fig. S1) and the microhabitat sampled, i.e. root or rhizosphere (Fig. S2). These results reinforce the preponderant role of soil type in the assembly of the root microbiome (Peiffer *et al*., 2013). Also, the higher diversity in the

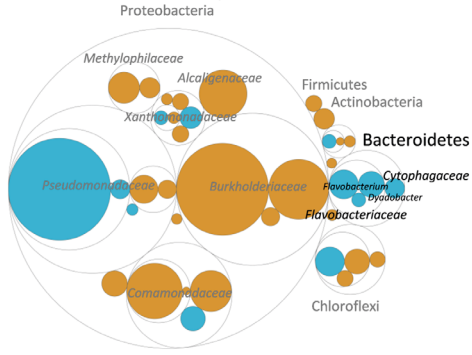
rhizosphere as compared to the endosphere (Fig. S2) is in accordance with previous reports (Edwards *et al*., 2015). Subsequent pairwise comparisons showed that, for each plant species, the Bacteroidetes were consistently enriched in the root or rhizosphere of wild relatives and a comparable difference was observed between *Cardamine hirsuta* and *Arabidopsis halleri* (moderated t-tests; $P<0.05$, BH corrected) (Fig. 1a). For the ancestor of sugar beet, *Beta vulgaris* ssp. *maritima*, we also observed a higher prevalence of Bacteroidetes taxa as compared to modern sugar beet, although this difference could not be analyzed statistically as the replicate samples in that study (Zachow *et al*., 2014) were pooled. Next to the Bacteroidetes, we observed a higher relative abundance of some other bacterial families on roots of wild relatives of the different plant species. In common bean, Planctomycetes, Verrucomicrobia and Acidobacteria together with some Proteobacteria families were also more abundant on roots of the wild accession. For wild barley, a few Proteobacteria families were enriched as well as two Firmicutes families. For wild lettuce and *Cardamine hirsuta,* also several Proteobacteria families were enriched. Overall, Proteobacteria and Actinobacteria were consistently enriched on roots of the modern counterpart, while Bacteroidetes was found almost exclusively enriched on roots of the wild relatives irrespective of the plant species and study. The phylum Bacteroidetes has also been found as a prevalent and abundant member in the rhizosphere of several other wild plant species (Aleklett *et al*., 2015; Shi *et al*., 2015).

a

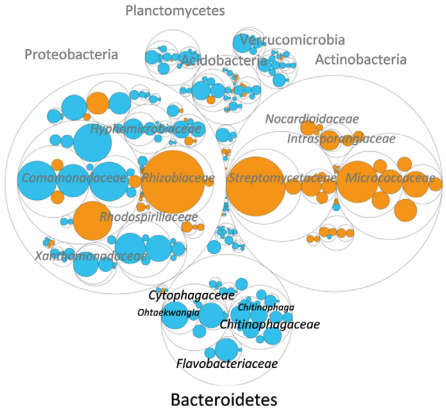
i) *Hordeum vulgare* ssp. *spontaneum* vs *Hordeum vulgare* ssp. *vulgare*



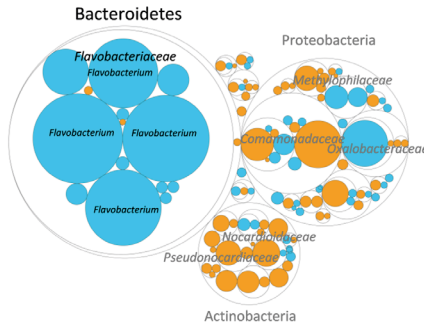
ii) *Lactuca serriola* vs *Lactuca sativa* ssp. *capitata*



iii) *Phaseolus vulgaris* (G22304-wild) vs *Phaseolus vulgaris* (G5773-modern)



iv) *Cardamine hirsuta* vs *Arabidopsis halleri*



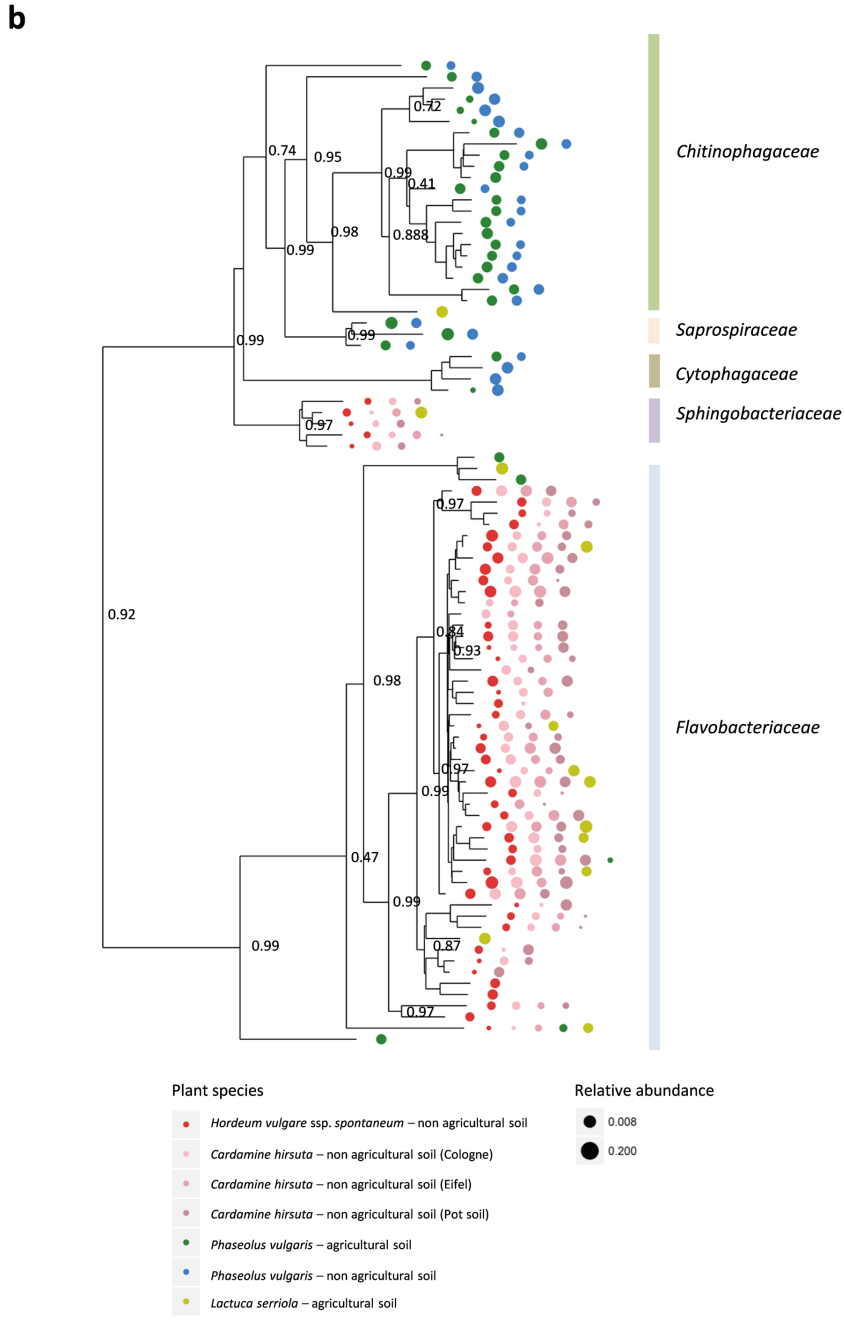


Fig. 1. Enrichment and taxonomic diversity of bacterial taxa in wild and domesticated plant species. a) Differential abundance of bacterial OTUs between wild plant accessions and their domesticated counterparts. Presented here are selected pairwise comparisons between **i)** wild barley (*Hordeum vulgare* ssp. *spontaneum*) and modern barley (*Hordeum vulgare* ssp. *vulgare*); **ii)** wild lettuce (*Lactuca serriola*)

and cultivated lettuce (*Lactuca sativa* ssp. *capitata*); **iii)** wild and modern accessions of common bean (*Phaseolus vulgaris*), and **iv)** *Cardamine hirsuta* and *Arabidopsis halleri.* Each comparison was made using a zero-inflated Gaussian distribution mixture model followed by moderated t-test and a Bayesian approach. Only OTUs significantly enriched in one of the two accessions are shown (FDR<0.05). The largest circles represent Phylum level and the inner circles represent Class and Family level. The color of the circles represents the OTUs enriched in the rhizosphere/roots of wild relatives (cyan) or of modern crop plants (orange), with the assigned genus in italics. The size of the circle is the mean read relative abundance of the differentially abundant OTU. b) Phylogenetic tree of bacterial members of the Bacteroidetes phylum associated with different wild plant species. The Bacteroidetes taxa were selected from microbiome data of wild plant species to construct the phylogenetic tree. The size of the circles corresponds to the relative abundance for each Bacteroidetes taxa. Only data with a relative abundance higher than 0.1% is depicted in the tree. Each abundance data is the average of at least 3 samples per plant species and site.

Our analysis further revealed that the extent of the Bacteroidetes enrichment on roots of wild plant relatives exhibits plant species-specific signatures. For example, approximately 50% of the bacterial species differentially enriched on roots of wild barley belonged to the Bacteroidetes, while for *Cardamine hirsuta,* wild lettuce and wild common bean the Bacteroidetes represented 33.3%, 24.5% and 18.9%, respectively, of the root-associated bacterial community. Subsequent phylogenetic analysis of the Bacteroidetes that were more abundant (>0.1%) on wild relatives showed two main clusters: one composed mainly of members of the *Chitinophagaceae* family and the other of members of the *Flavobacteriaceae* family (Fig. 1b). The family *Flavobacteriaceae* was represented by a high diversity in *Cardamine hirsuta* and wild barley, whereas *Chitinophagaceae* and *Cytophagaceae* families were predominant in the root microbiome of wild relatives of common bean (Fig. 1b). Collectively, these results indicate that plant domestication resulted in a similar overall taxonomic shift in the prokaryotic root microbiome with a reduced abundance of the Bacteroidetes phylum on modern accessions and a concomitant increase of members of the Actinobacteria and Proteobacteria (Fig. 2). At higher taxonomic levels, we observed that the plant species-specific effects observed on Bacteroidetes families may be probably due to differences in the physicochemical

characteristics of the diverse soils used in these independent studies, such as divergent pH values and the organic carbon content (Table S1).

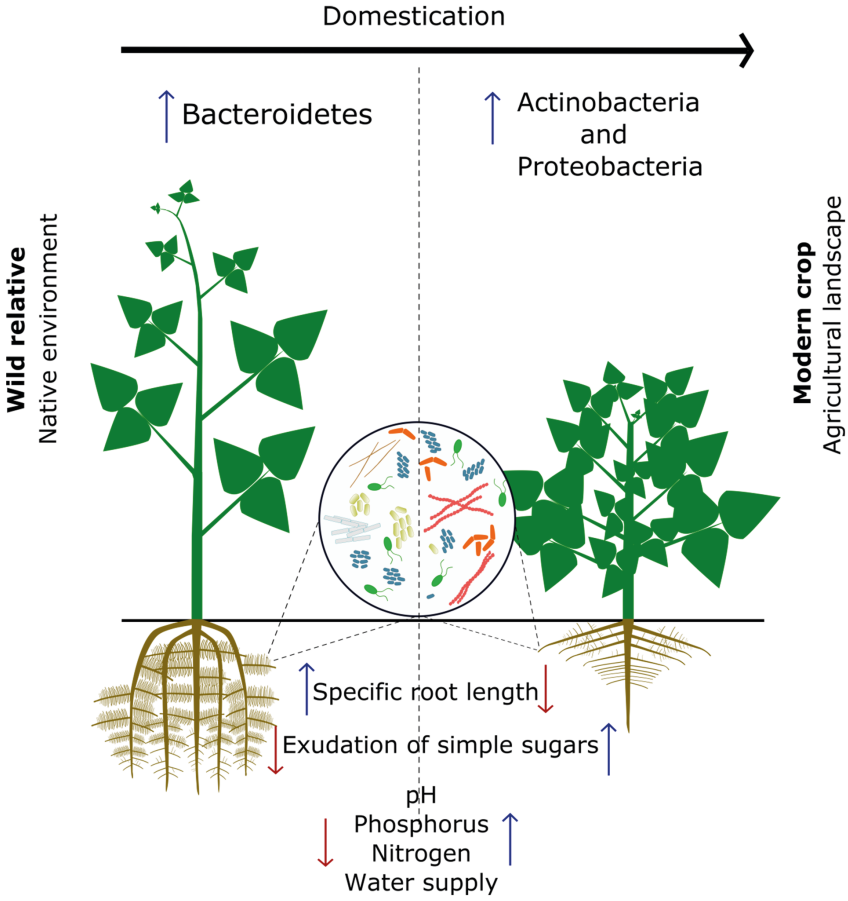


Fig. 2. Impact of domestication on soil management, plant phenotype, plant physiology and rhizobacterial diversity. In this hypothetical schematic representation, the root morphology of the wild relative substantially differs from that of the modern counterpart. Readily available macronutrients and water associated with agricultural management led to shallower roots in modern crop cultivars as compared to roots of wild relatives, which are rooting deeper with conspicuous lateral roots. Domesticated crop plants presumably also exude more 'simple' sugars than their wild relatives. The impact of the domestication process on rhizobacterial community composition is reflected in a decrease in Bacteroidetes abundance on modern crop plants, while the abundances of the Actinobacteria and Proteobacteria are increased.

Firstly, it is important to emphasize that in our analyses the same computational pipeline was used adopting a rarefaction of the OTU table to the same sequencing depth. However, the approach used in this study cannot address all biases associated with this type of meta-analysis. Differences in soil types, sampling strategies, nucleic acid extraction protocols and sequencing techniques between the different studies may have affected the reach of our meta-analysis and the interpretation of the results. Nevertheless, it is noteworthy that despite all these constraints we found similar and consistent differences between the prokaryotic composition of the root/rhizosphere microbiome of wild and domesticated plant species with a significantly higher abundance of Bacteroidetes on/in roots of wild plant relatives. Why Bacteroidetes are relatively more abundant in the root and rhizosphere compartments of wild relatives of various crop plant species is yet unknown. They are recognized for their ability to degrade complex biopolymers, a trait associated with a diverse set of carbohydrate processing enzymes (Thomas *et al*., 2011; Berlemont *et al*., 2015). Hence, their prevalence in the root compartments of wild plant species may be a phylogenetic signal associated with the presence of complex biopolymers in their root exudates (Fig. 2). Plant root exudates can have a major impact on the structure and functioning of microbial communities in soil environments (Bais *et al*., 2006; Micallef *et al*., 2009a). A recent study on mutants of poplar trees, silenced in the cinnamyl-Co reductase (*CCR*) gene of the monolignol-specific lignin pathways, showed significant effects on the density and composition of culturable rhizosphere and endosphere bacteria, microbiome shifts that were proposed to be mediated, at least in part, by changes in extractable plant phenolic compounds such as ferulic acid (Beckers *et al*., 2016). In this context it is worth noting that one of the most common domestication syndrome traits is related with changes in the type and amount of secondary metabolites, such as the loss of specific compounds that are toxic for humans or livestock or the reduction of flavonoid content in the leaves (Gepts, 2004; Meyer *et al*., 2012; Chacón-Fuentes *et al*., 2017). To

date, however, very little is known about qualitative and quantitative differences between root exudation profiles of crop plants and their wild relatives. For wheat, it has been shown that a modern wheat variety exuded three to five times more 'simple' sugars (mainly fructose, glucose and maltose) than an ancient wheat cultivar under stress conditions, a feature that might be related with a lower capacity of the modern wheat cultivar to control sugar exudation (Shaposhnikov *et al*., 2016). Whether the higher levels of these 'easy-digestible' sugars are also the case for other plant species and may contribute to a competitive advantage and a concomitant higher abundance of Proteobacteria and Actinobacteria on roots of modern crop cultivars remains to be addressed.

Also differences in root architecture between crop plants and their wild relatives may impact root microbiome assembly. More specifically, the prevalence of Bacteroidetes in the rhizosphere of wild bean correlated significantly with a higher specific root length (SRL, i.e. root length per unit of root dry mass) and a lower root density (Pérez-Jaramillo *et al*., 2017). A high SRL has been associated with a higher efficiency of water search and uptake for the plant and is considered a strategy to acquire nutrients in low-fertile soils (Comas *et al*., 2013; Kramer-Walter *et al*., 2016). Along with changes in plant genotype and phenotype, the domestication process also involves changes in the environment and the concomitant need of management practices, such as the use of chemical pesticides and fertilizers, to sustain growth and health of the crop plants (Pérez-Jaramillo *et al*., 2016). Therefore, altered root morphology traits (Fig. 2) as well as changes in plant physiology and root exudation may have contributed to the observed and consistent shifts in the prokaryotic root microbiomes between wild plant relatives and their domesticated counterpart. This hypothesis needs to be validated by experiments where morphological and physiological traits, in particular root architecture and

exudation profiles, of wild relatives of crop plants are assessed in agricultural soils as well as in soils from their centres of origin and diversification.

Whether a higher relative abundance of Bacteroidetes affects plant growth and health as was shown for growth (i.e. obesity) and health of humans (Ley *et al*., 2006; Arrieta *et al*., 2014; Liu *et al*, 2017) is not known to date. Some studies suggested that representatives of this phylum can affect plant growth and health. In particular strains of the genus *Flavobacterium* have been associated with plant growth promotion and disease protection (Kolton *et al*., 2012). For the legume plant *Trifolium pratense*, however, *Flavobacterium* led to impaired shoot growth (Hartman *et al*., 2017). For the genus *Chryseobacterium,* disease protective effects have been described (Yin *et al*., 2013), but effects on plant growth and health by most other Bacteroidetes, including members of the *Chitinophagaceae* and *Cytophagaceae* families detected here, remain to be discovered. Establishing a phenotypically and genomically diverse and well-characterized collection of Bacteroidetes species from multiple wild plant relatives followed by controlled bioassays to test the effects of individual species/strains and consortia on plant growth and health under diverse environmental conditions will shed more light on their functional importance for the growth and survival of wild plant species in their native, environmentally harsh habitats. Understanding the functional importance of these 'missing plant microbes' can be highly instrumental in plant breeding programs and for improving our future crop production systems in a changing environment.

Acknowledgments

We thank the authors of the studies used in the meta-analysis for kindly providing the raw data and the metadata files. JEP-J was financially supported by the Department of Science, Technology and Innovation of Colombia—COLCIENCIAS through the doctoral

grant 568-2012-15517825. JMR and VJC were supported by the Dutch NWO-TTW Perspectief program ‘Back to the Roots’.

Supplementary materials

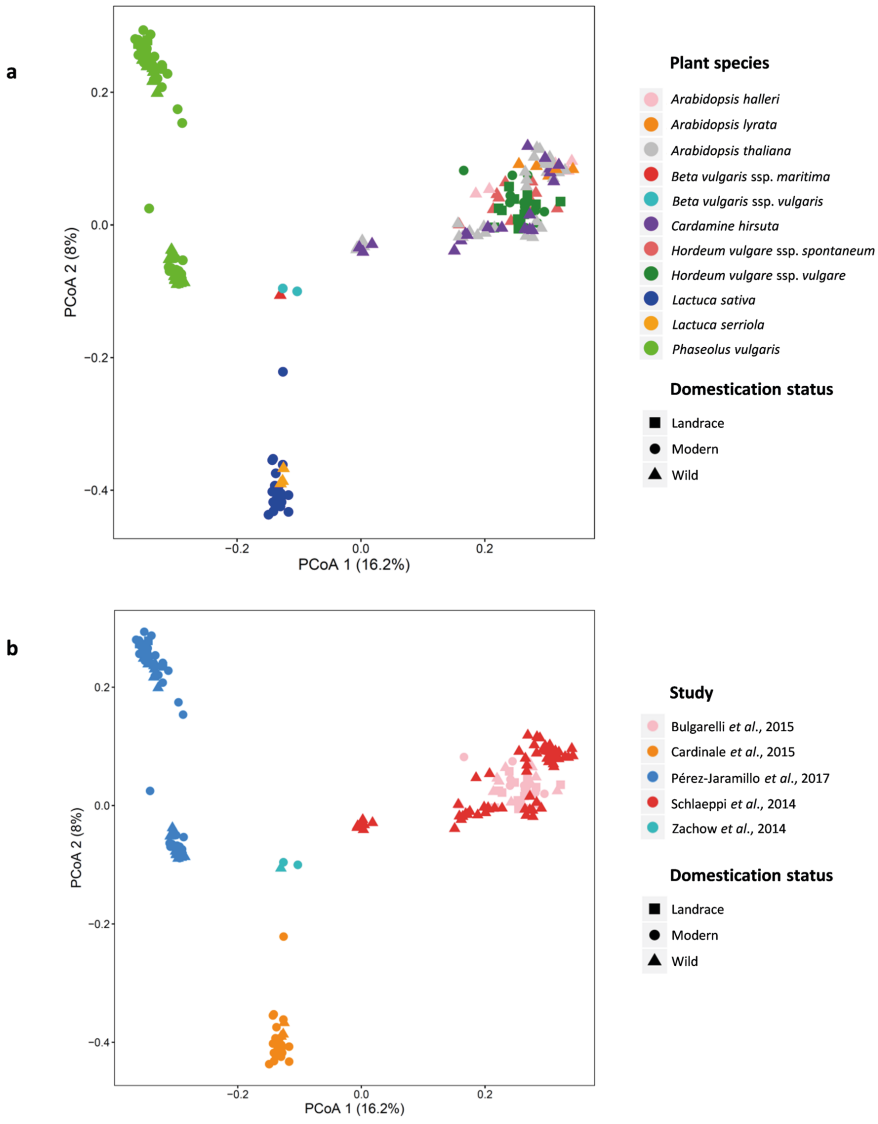


Fig. S1. Rhizosphere bacterial community composition across studies of wild, landrace and modern plants. Principal Coordinate Analysis (PCoA) of Bray-Curtis dissimilarities of 16S rRNA data. **a)** PCoA with samples colored by plant species. b) PCoA with samples colored by study. The source of the data (study) was the main explaining variable as assessed by PERMANOVA (29.1%; P<0.005).

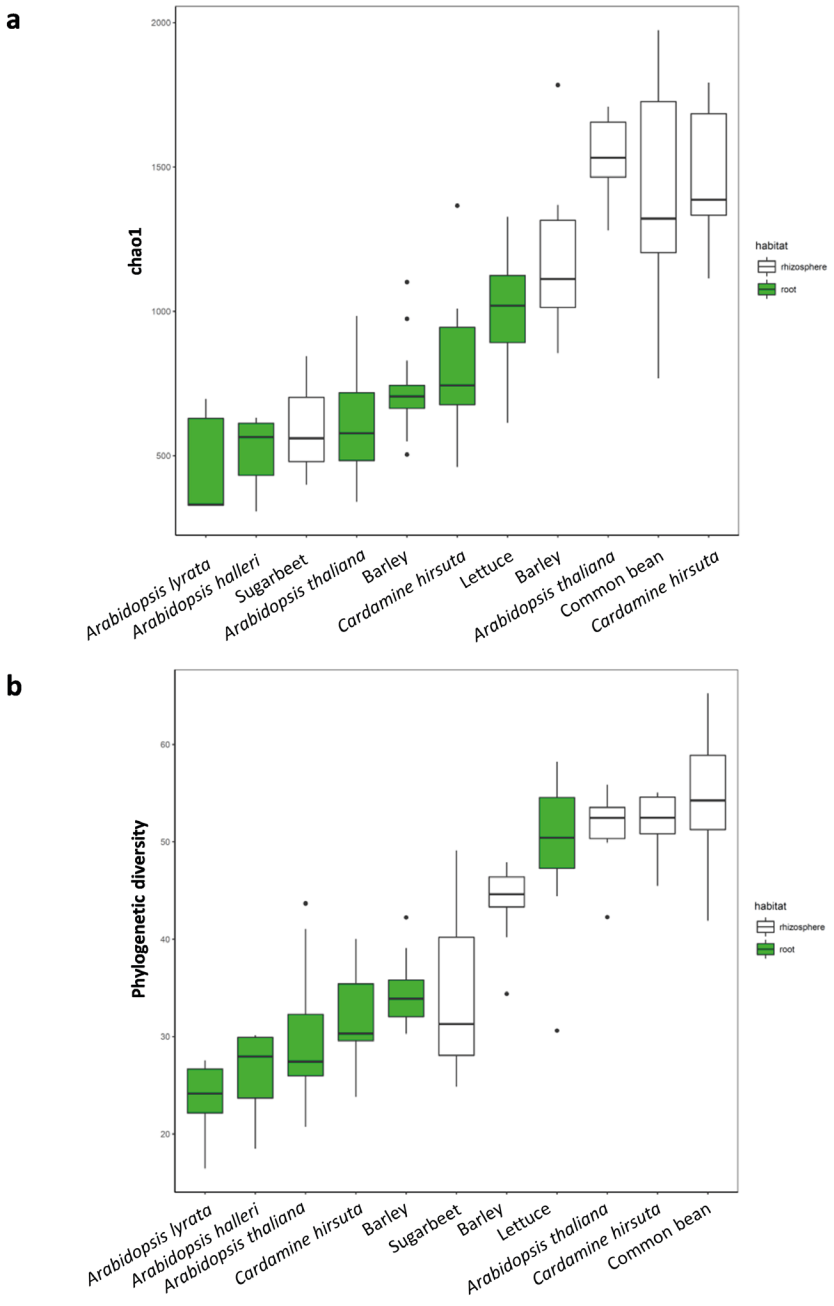


Fig. S2. α-diversity of 16S sequence data of wild, landrace and modern plant species for rhizosphere and roots. (a) Chao1 and **(b)** Phylogenetic diversity for all the plants included in the meta-analysis. Bacterial diversity on/in the roots is less than in the rhizosphere.

Table S1. General information of the datasets used for the meta-analysis

Study	Sequencing Method	Repository	Project code	16S primer pair	Plant species and subspecies	Domestication status	DOI
Zachow *et al.,* 2014	454	NCBI*	PRJNA233435	V4-V5	*Beta vulgaris* ssp. *vulgaris*	Domesticated beet	10.3389/fmicb.2014.00415
					Beta vulgaris ssp. *maritima*	Wild relative	
Schlaeppi *et al.* , 2014	454	ENA**	PRJEB5058	V5-V6-V7	*Cardamine hirsuta*	Wild plant	10.1073/pnas.1321597111
					Arabidopsis halleri	Wild plant	
					Arabidopsis lyrata	Wild plant	
					Arabidopsis thaliana	Wild plant	
Bulgarelli *et al.,* 2015	454	ENA	PRJEB5860	V5-V6-V7	*Hordeum vulgare* ssp. *spontaneum*	Wild relative	10.1016/j.chom.2015.01.011
					Hordeum vulgare ssp. *vulgare*	Domesticated barley	
Cardinale *et al.,* 2015	454	ENA	PRJEB5101	V4	*Lactuca serriola*	Wild relative	10.1111/1462-2920.12686
					Lactuca sativa ssp. *capitata*	Domesticated lettuce	
					Lactuca sativa ssp. *crispa*	Domesticated lettuce	
					Lactuca sativa ssp. *longifolia*	Domesticated lettuce	
					Lactuca sativa ssp. *augustana*	Domesticated lettuce	
Leff *et al.,* 2017	MiSeq	NCBI	SRP075934	V4	*Helianthus annuus*	Wild, landraces and modern accessions included ***	10.1111/nph.14323
Pérez-Jaramillo *et al.,* 2017	MiSeq	ENA	PRJEB19467	V3-V4	*Phaseolus vulgaris*	Wild, landraces and modern accessions included ***	10.1038/ismej.2017.85

*National Center for Biotechnology Information of the United States of America

**European Nucleotide Archive

***Wild; plants that have a ready ability to grow freely in natural ecosystems, with strong dispersal mechanisms; Landrace: locally developed crop varieties by farmers within their own agricultural, horticultural or agri-silvicultural systems; Modern accessions or modern improved varieties are the result of plant breeding in the pursue of higher yields, better quality and more stable production, typically grown on heavily managed agricultural settings.

Table S2. Physicochemical characteristics of the soils used in the studies included in the meta-analysis.

Study	Soil type used for plant growth						Type of experiment
	Texture (%)			Classification	pH	Organic C (%)	
	Clay	Silt	Sand				
Zachow *et al.,* 2014	NA*	NA	NA	Clay	9,5	NA	Field/Pot
Schlaeppi *et al.* , 2014	13,4	37,3	49,3	Sandy Loam	6,95	4,0	Pot trial
Bulgarelli *et al.,* 2015	4,2	4,2	91,6	Sand	7,12	1,0	Pot trial
Cardinale *et al.,* 2015	NA	NA	NA	NA	NA	NA	Field
Leff *et al.,* 2017	NA	NA	NA	Sandy Loam	NA	NA	Field
Pérez-Jaramillo *et al.,* 2017	8	30	62	Clay Loam	5,8	17.9**	Pot trial
*Not available							
*Organic Matter							

Chapter 6

The spermosphere microbiome of wild and domesticated common bean (*Phaseolus vulgaris*)

Juan E. Pérez-Jaramillo, Víctor J. Carrión and Jos M. Raaijmakers

Abstract

Plants have a significant influence on the diversity and activity of soil microbial communities. Already during imbibition and germination, plant seeds release exudates that promote microbial growth. Although this so-called spermosphere effect is widely studied, little is known about the impact of plant domestication on spermosphere microbiome assembly. Here, we sampled the spermosphere from germinating seeds of seven wild and modern accessions of common bean (*Phaseolus vulgaris*) and assessed the bacterial community composition by 16S rRNA sequencing. Compared to bulk soil we observed a decrease in α-diversity of the bacterial community in the spermosphere of all bean accessions. Also for the β-diversity, a significant difference was observed between the bean spermosphere and bulk soil. Proteobacteria, Actinobacteria and Firmicutes were enriched in the spermosphere. Small but significant differences in β-diversity were detected between the spermosphere microbiomes of wild and modern bean accessions, suggesting a genotype-dependent effect on microbiome assembly already at this early plant developmental stage. Comparative seed exudate profiling further showed that the spermosphere of modern bean accession M5 contained higher levels of glutamate and glutamine as compared to that of wild accession A2. Subsequent in vitro assays indicated that these single amino acids did not enrich for the specific bacterial taxa found at higher abundance in the spermosphere of modern bean accessions. Collectively, these results suggest that domestication affects spermosphere microbiome assembly of common bean and that more complex mixtures of yet unknown exudate constituents are driving spermosphere microbiome assembly.

Keywords: spermosphere, microbiome, root exudates, glutamate, Actinobacteria

Introduction

Early in plant development, during seed imbibition and germination, exudates are released in the surrounding soil creating a narrow zone with enhanced microbial activity known as the spermosphere (Nelson *et al*., 2004; Schiltz *et al*., 2015; Lemanceau *et al*., 2017). In the spermosphere, a broad diversity of compounds can be found such as sugars, amino acids, phenolic compounds and volatiles (Nelson, 1990; 2018). The composition of exudates in the spermosphere is determined by abiotic factors such as temperature, oxygen tension, soil moisture, seed coat integrity and genetic traits (Short and Lacy, 1976; Schlub and Schmitthenner, 1978; Nelson, 2018). Different plant species and even cultivars from the same species can differ in seed exudation profiles (Kageyama and Nelson, 2003; Singh and Mehrotra, 1980). Consequently, seed exudation may actively influence the spermosphere microbiome composition and activity in a plant genotype-dependent manner. For instance, two contrasting tomato recombinant inbred lines (RIL) populations derived from an interspecific cross of cultivated tomato and a wild tomato relative, differed in the proportion of indigenous fluorescent *Pseudomonas* species in the spermosphere (Simon *et al*., 2001). Similarly, a host genotype-dependent effect was found for the fungal community composition in the spermosphere (Barret *et al*., 2015). To date, however, little is known about the diversity, succession, and activities of indigenous microbial communities in the spermosphere and whether the process of plant domestication influenced spermosphere microbiome assembly. Finally, it is unknown how this short-lived, dynamic spermosphere community affects plant growth, development and health.

The objectives of this study were to i) study the impact of the spermosphere of common bean (*Phaseolus vulgaris*) on bacterial diversity and abundance, and to ii) decipher differences in spermosphere microbiome assembly between modern bean cultivars and wild relatives. To this end, we selected wild and modern accessions of common bean for

which significant differences in the rhizobacterial community assembly were documented earlier (Pérez-Jaramillo *et al*., 2017). Seeds from the common bean accessions were sown in agricultural soil collected from the Colombian highlands and spermosphere soil samples were collected at 48h following seed imbibition and germination. The bacterial communities were characterized through 16S rRNA amplicon sequencing and cultivation-dependent methods. We also conducted comparative chemical profiling of seed exudates of a wild and a modern bean accession and assessed the effect of individual differential amino acids on soil bacterial community composition.

Materials and Methods

Soil and plant material

The agricultural soil used for the assays was collected in a common bean producing farm in the rural area of the municipality of El Carmen de Viboral, Colombia, as described in Pérez-Jaramillo *et al.* (2017). Briefly, the soil was air dried, sieved (2-mm mesh) and stored at 4°C for further use. The seeds of two wild accessions (A1 and A2) and five modern cultivars of common bean (M1 to M5) were selected for this study. The plant material was kindly provided by the Genetic Resources Program at the International Centre for Tropical Agriculture—CIAT—in Palmira, Colombia. A complete description of these bean accessions can be found in Pérez-Jaramillo *et al.* (2017)

Sampling of soil from the spermosphere compartment

Seeds of wild and modern accessions of common bean without cracks or other visible damages were selected. Seed-surface disinfection was applied to the seeds through immersion during 3 minutes in sodium hypochlorite 0.5% (v/v), followed by four washes with ample sterile distilled water. After disinfection, the seeds were air dried in a laminar flow cabinet. For each bean accession, surface-sterilized seeds were sown in 12-well

plates with 3g of the agricultural soil and one seed per well. For each bean accession, 6 replicates were used. Wells with soil but without seeds served as the bulk soil control. After sowing the seeds at a depth of 0.5 cm, 1mL of sterile distilled water was added to each well. This initial watering step initiated seed imbibition. The plates were kept at 25 °C in the dark during 48-72 hours, depending on the differences in germination rates between the seven bean accessions. Once the radicle protruded, the seeds were harvested and replaced by a new seed of the same accession. This first imbibition and germination cycle was done to pre-condition the soil with seed exudates in order to activate the soil microbial communities. In the second cycle, a new seed was sown in the pre-conditioned soil and 1 mL of sterile distilled water was added. The plates were kept at 25 °C in the dark during 48-72 hours. Once the radicle protruded, a round hole puncher (1,5 cm diameter) was used to extract the germinating seed with adhering spermosphere soil from each well. The seed with spermosphere soil was transferred to a 15 mL tube containing 5 ml of LifeGuard Soil Preservation Solution (MoBio Laboratories). The 15mL tubes were thoroughly vortexed at maximum speed during 15 min, and the solution was kept at -20°C for further use.

DNA isolation, sequencing and preparation of the OTU table

For each bean accession and the control (bulk soil), four replicates were used for DNA extraction with MoBio kits (MoBio Laboratories, Carlsbad, CA, USA) as previously described (Pérez-Jaramillo *et al*., 2017). Subsequently, the V3-V4 region of the 16S rRNA was amplified and sequenced by Illumina Myseq (Illumina, 2013). The sequence data analysis was performed by the Hydra pipeline implemented in Snakemake (Köster and Rahmann, 2012) and the obtained OTU table was filtered using QIIME 1 (1.9.1) custom scripts (Kuczynski *et al*., 2012). The Bacteria domain was extracted using the command *split_otu_table_by_taxonomy.py* and singletons, doubletons, chloroplast and

mitochondria sequences were discarded with the command *filter_otus_from_otu_table.py*, obtaining a filtered OTU table for further analysis.

Bacterial diversity, differential abundance and core microbiome analysis

To calculate α-diversity, we rarefied (Brewer and Williamson, 1994; Gotelli and Colwell, 2001) the OTU table up to 45,360 reads (i.e. lowest sequencing depth obtained from a sample) per sample for all spermosphere and bulk soil samples using the script *alpha_rarefaction.py* retrieved from QIIME 1 (1.9.1). Chao1 and Shannon diversity indices were calculated using the *alpha_diversity.py* command and the data obtained was tested for normality and homoscedasticity. To compare α-diversity values by bean accession, the non-parametric Kruskal-Wallis test was applied using R (R Core Team, 2015). To calculate β diversity, the unrarefied OTU table was normalized using a cumulative-sum scaling (CSS) method (Paulson *et al*., 2013). Bray-Curtis dissimilarities followed by Constrained Analysis of Principal Coordinates were calculated with the Phyloseq package (McMurdie and Holmes, 2013) (v.1.22.3). The nonparametric *adonis* test was used to assess the percentage of variation explained by the habitat (bulk soil, spermosphere) and domestication status (wild, modern) along with its statistical significance using Vegan (Oksanen *et al*., 2016) (v.2.4-6), all performed in R. The QIIME command *compute_core_microbiome.py* was applied to the rarefied OTU table in order to obtain a list of OTUs observed in the spermosphere of all bean accessions. Core OTUs with a relative abundance <0.5% and unclassified taxa were not included in the figure. Donut plots were built using the R package ggplot2 (Wickham, 2009) (v. 2_2.2.1). To assess the spermosphere effect on microbiome composition, the rarefied OTU table was analyzed with the STAMP software (v.2.1.3) (Parks *et al*., 2014). Welch's t-tests followed by BH corrections were performed at phylum, family and genus level between spermosphere samples grouped by the domestication status. Bacterial phyla, families and

genera significantly overrepresented ($P < 0.05$) in the spermosphere or in the bulk soil microbiome are depicted using extended error bars. Finally, Welch's t-tests followed by BH corrections were performed to assess if bacterial genera were differentially abundant in the spermosphere of wild and modern bean.

Culture-dependent characterization of the bacterial diversity in the spermosphere

After extracting the spermosphere soil, an aliquot of 500 mg was suspended in 0.1 M phosphate buffer. After 5min of rigorous vortexing, serial dilutions of the soil suspension were made and 100 µL aliquots of 10^{-5}, 10^{-6}, and 10^{-7} dilutions were plated onto $1/10^{th}$ strength R2A agar plates. The plates were incubated at 25 ºC for 1 week and isolation of individual colonies was performed based on morphology, colour and size. A total of 631 isolates was purified and kept in R2A broth with 40% (v/v) of glycerol at -80 ºC for further use. For identification purposes, the collection was replicated and the V1-V4 16S rRNA region was sequenced at BaseClear (Leiden, Netherlands). Multiple alignment of sequences larger than 400 bp was performed using Muscle v3.7 (Edgar, 2004) with the default parameters to create an approximate maximum-likelihood tree in FastTree v2.1.8. with default parameters (Price *et al*., 2009). The tree was visualized with the online application iTOL (Letunic and Bork, 2016).

Exudate extraction from seeds

Seeds without cracks or visible damages in the seed coat were selected for seed exudate extractions. Due to seeds limitations of several accessions, only wild accession A2 and modern accession M5 were used for exudate extraction. Firstly, seeds were disinfected as described above. From the last wash 20 µL of water were plated in TSA and PDA petri dishes and incubated 48h to confirm disinfection. One gram of disinfected seeds (i.e. approximately 6 seeds) was transferred to sterile petri dishes containing sterile filter paper

(Whatman no. 1). For each bean accession, 4 petri dishes with 1 g of seeds were used. The filter paper was wetted with 200 µL sterile distilled water to saturation point and incubated at 24ºC for 24 hours. After this time, the seeds were transferred to a new sterile plate with sterile filter paper (Whatman no. 1), and wetted with 200 µL sterile distilled water to saturation point and incubated at 24ºC for another 24 hours. The filter papers were processed individually to harvest the exudates. 1 mL of a Methanol:Water solution (70:30) was added to each of the filter papers and incubated during 1 hour on an orbital shaker (100 rpm). The resulting extract was filtered (2 µm) and dried under a flow of Nitrogen gas. Once the methanol was evaporated, the samples were freeze-dried for three days under vacuum (−80 °C collector temperature; Labconco Free Zone 12 L Freeze Dry System, USA). Dried samples were re-suspended in 200 µL of MilliQ water, filtered through a 13-mm-diameter 0.2 µM PTFE syringe filter and poured in glass vials for chemical analyses. Organic acids were analysed by HPLC (Ultimate 3000, ThermoFisher Scientific, USA) equiped with a UV diode array detector. The sugar analyses were performed by HPLC with electro-chemical detection, whereas amino acid analyses were performed by LC-MS/MS (Agilent, USA). Differences in the amount of exudates between the wild and the modern accessions were determined by the Wilcoxon signed-rank test ($P < 0.05$) in R.

Effect of glutamate on bacterial community assembly

The agricultural soil used for spermosphere microbiome analyses (described above) was used to test the effect of glutamate on bacterial community assembly. Four treatments were evaluated as follows: T1 consisted of M9 minimal media (12.8 g $Na_2HPO_4 \cdot 7H_2O$, 3 g KH_2PO_4, 0.5 g NaCl, 1 g NH_4Cl in 1L of distilled water); T2 consisted of M9 minimal media, 10 mM of glutamate, 1 mM of $MgSO_4$, and glucose 20%; T3 consisted of M9 minimal media, 1 mM of $MgSO_4$, and glucose 20%; and T4 consisted of M9 minimal

media, 10 mM of cysteine, 1 mM of $MgSO_4$, and glucose 20%. Treatment T1 was considered the baseline control, because only had minimal medium without glucose or any type of amino acid. Treatment T2 is the glutamate treatment. Treatment T3 is the glucose control, in order to observe the impact on soil bacterial communities of the glucose alone. Finally, T4 is the amino acid control treatment, which consisted of cysteine, glucose and minimal media. Cysteine was included as a control amino acid due to the low quantity observed in seed exudation profiles, and the assumption that its impact on bacterial communities would be not significant. The experimental unit consisted of a 50 mL conical flask, and to each flask 3 g of the agricultural soil were added. As 4 replicates per treatment were used, in total 16 conical flasks were used for the experiment. At time 0, the respective treatments were applied to the flasks and these were kept in an orbital shaker at 150rpm and 25 ˚C. After 24 hours, the content of the flasks was transferred to 50 mL tubes carefully labelled and centrifuged at 3,500 rpm for 5 min. After centrifugation, the supernatant was removed and the treatments were re-applied, i.e, to each 50 mL tube with soil a new addition of the respective treatment solution was applied, and after vigorous mixing, the soil solutions were transferred back to their respective conical flasks for a new incubation period of 24 h. This enrichment procedure was repeated five times followed by soil DNA extraction, amplification and sequencing of the V3-V4 region of the 16S rRNA, sequence analysis and OTU table preparation, the same procedures as described above for the spermosphere microbiome analyses. The OTU table was rarefied up to 11,790 reads and bar plots with the relative abundance of phyla and families were built based on the normalized OTU table. Also a Bray-Curtis dissimilarity matrix was calculated and used for Principal Coordinate analysis with Phyloseq. The nonparametric *adonis* test was used to assess the percentage of variation explained by the glutamate treatment along with its statistical significance using Vegan package, all performed in R.

Results

Spermosphere microbiome assembly is plant genotype specific

16S rRNA sequencing of the spermosphere and bulk soil samples yielded approximately 2.2 million reads after quality filtering, representing 8,263 bacterial operational taxonomic units (OTUs) at 97% sequence similarity. For the α-diversity, we observed significantly lower chao1 and Shannon indices for the spermosphere as compared to the bulk soil (Fig. 1a and b). Regarding the β-diversity, Bray-Curtis metrics and Constrained Principal Coordinate Analysis (CAP) revealed a significant effect of the spermosphere and the domestication status of the bean accessions (Fig. 1c and d). The structure of the bacterial communities differed between the spermosphere and the bulk soil samples: the percentage of variability explained by the habitat (spermosphere versus bulk soil) was 28.2% (PERMANOVA, P=0.001) (Fig. 1c). Furthermore, a subtle but statistically significant difference was found between the spermosphere microbiome composition of wild and modern bean accessions with 5.2% of the total variation explained by the domestication status (wild versus modern; PERMANOVA, P=0.001) (Fig. 1d).

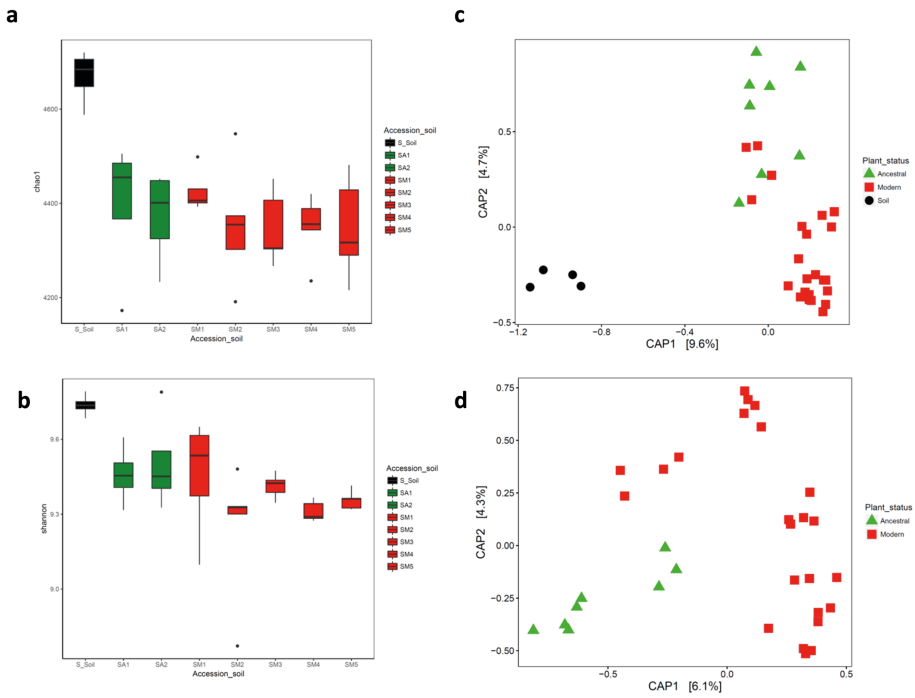


Fig. 1. Bacterial diversity and community structure of spermosphere and bulk soil samples. (a) chao 1 and (b) Shannon were calculated for all bean accessions and the bulk soils samples. The data was rarefied up to 45360 counts per sample. Additionally, Constrained Principal Coordinate Analysis (CAP) of 16S rRNA diversity in the spermosphere of the seven common bean accessions were performed, with (c) and without (d) 16S rRNA diversity in the bulk soil, respectively. Green color was assigned to ancestral accessions, red color was assigned to modern accessions and black color was assigned to bulk soil samples.

The spermosphere significantly impacts bacterial abundances

From the total of 8,263 clustered OTUs, we found that 1,252 OTUs comprised the core spermosphere microbiome of the seven bean accessions. When classified up to genus level, the core microbiome of 1,252 OTUs represented 17.8% of the total number of OTUs and the overall majority (89%) of all the sequence reads detected. The genus *Sphingomonas* was the most abundant genus in the spermosphere core microbiome

(6.6%), followed by *Pseudarthrobacter* (4.8%), *Burkholderia-Paraburkholderia* (2.5%), *Bradyrhizobium* (2.5%) and *Bacillus* (2.4%) (Fig. 2). Another genus found in the core spermosphere microbiome of common bean was *Rhizobium*, with a relative abundance of 1.6%.

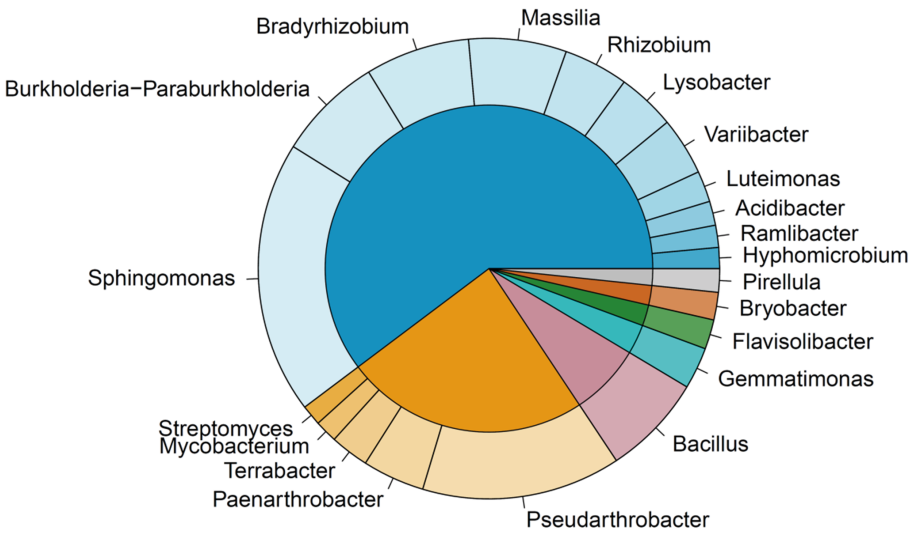


Fig. 2. Core spermosphere microbiome of common bean. The different portions within the inner pie chart represent the bacterial phyla that are part of the spermosphere core microbiome. The outer donut plot represents the genera that are part of the core and each genus assigned to the phylum they belong to. The size of the different pie and donut portions represents the contribution of each phylum/genus to the total relative abundance.

To determine the selective effect of the spermosphere on microbiome composition, the bacterial abundances of spermosphere samples from wild and modern common bean accessions were compared first to bulk soil samples. At Phylum level, a significant increase was observed for Proteobacteria, Firmicutes, Bacteroidetes and Saccharibacteria in the spermosphere of both wild and modern common bean accessions as compared to bulk soil (Fig. S1). Conversely, bacterial phyla such as Chloroflexi, Acidobacteria, Verrucomicrobia and Planctomycetes were significantly more abundant in bulk soil (Fig.

S1). At family level, *Bacillaceae*, *Oxalobacteraceae*, *Burkholderiaceae*, *Rhizobiaceae*, *Xanthomonadaceae*, *Comamonadaceae*, *Methylophilaceae* and *Paenibacillaceae* were more abundant in the spermosphere microbiome of wild bean accessions than in bulk soil (Fig. S2). The spermosphere microbiome of modern bean accessions was enriched with the same bacterial families as well as with *Micrococcaceae*, *Mycobacteriaceae*, *Streptomycetaceae*, *Nocardiaceae Sphingomonadaceae*, and *Sphingobacteriaceae* (Fig. S3). At genus level, *Bacillus, Burkholderia-Paraburkholderia, Rhizobium, Massilia*, *Lysobacter, Paenarthrobacter, Luteimonas, Ramlibacter* and *Paenibacillus* were enriched in the spermosphere of both wild and modern bean accessions as compared to bulk soil (Figs. S4 and S5). The spermosphere of modern accessions was also enriched with *Sphingomonas*, *Pseudarthrobacter, Paenarthrobacter, Mycobacterium, Arenimonas, and Streptomyces* (Fig. S5). Comparing abundances at bacterial family and genus levels between the spermosphere of wild and modern common bean accessions showed that the families *Streptomycetaceae, Mycobacteriaceae* and *Nocardiaceae* were enriched in the spermosphere of the modern common bean accessions (Fig. 3a). At genus level, *Paenarthrobacter*, *Streptomyces*, *Mycobacterium*, *Rhodococcus* and *Geodermatophilus*, all from the phylum Actinobacteria, were enriched in the spermosphere of modern common bean (Fig. 3b).

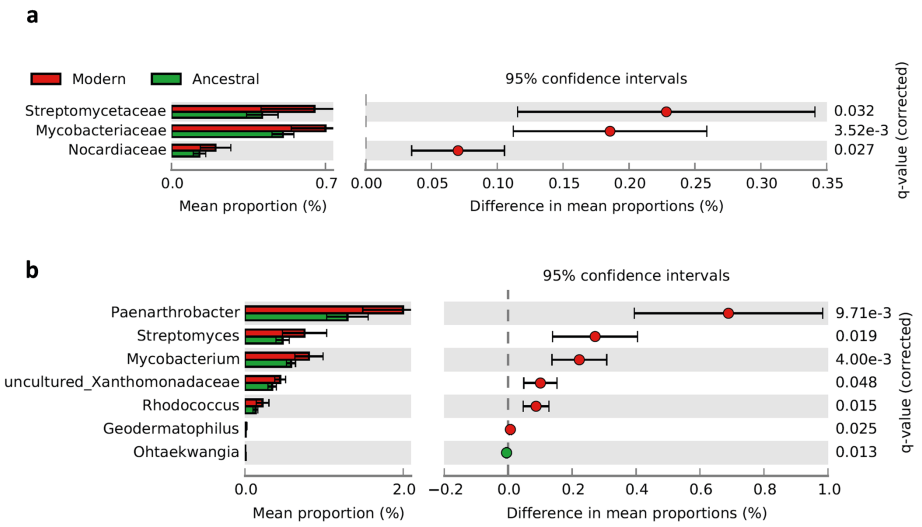


Fig. 3. Relative abundance of bacterial families and genera in the spermosphere of wild and modern bean accessions. Welch's t-tests followed by BH corrections were performed. Only families and genera differentially abundant are shown in the extended error bar plots ($P < 0{,}05$).

Bacterial isolation captured the majority of abundant core microbiome genera

The culture-dependent isolation and 16S rRNA characterization of bacteria from the bean spermosphere showed that from the core microbiome described above, only the Proteobacterial genera *Variibacter*, *Luteimonas*, *Acidibacter* and *Hyphomicrobium* were not isolated on the agar medium used together with a few genera from Acidobacteria, Planctomycetes and Gematimonadetes phyla. Amongst the 596 sequenced isolates, we found in total 67 genera belonging to the Actinobacteria, Proteobacteria, Firmicutes and Bacteroidetes phyla (Fig. S5). Actinobacteria was the most isolated phylum with a total of 306 isolates and 24 genera. *Streptomyces*, *Terrabacter, Arthrobacter, Nocardioides* and *Phycicoccus* were the most abundant genera amongst the Actinobacterial isolates with 66, 47, 24, 19 and 19 isolates obtained, respectively. The second most isolated phylum was Proteobacteria with a total of 185 isolates and 34 genera. *Lysobacter*, *Massilia*, *Burkholderia-Paraburkholderia*, *Rhizobium* and *Sphingomonas* were the most abundant genera within Proteobacteria, comprising 33, 26, 18, 16 and 11 isolates per

genus, respectively. From the Firmicutes, a total of 100 isolates were recovered belonging to 5 genera, with *Paenibacillus* (58 isolates) and *Bacillus* (41) representing the majority. Finally, from the phylum Bacteroidetes, 4 isolates were recovered and classified as *Dyadobacter*, *Flavisolibacter*, *Pedobacter* and *Terrimonas*. Additionally, when the V3-V4 region sequences from the 16S rRNA spermosphere dataset were matched with the sequences from the V1-V4 region from the isolates, we observed that more than 66.6% of the genera identified with the cultivation-independent approach as part of the core (relative abundance >0.5%) were successfully recovered with the cultivation-dependent approach (Table 1).

Seed exudate composition of wild and modern bean accessions

Seed exudates were collected after 24 and 48 hours incubation and, due to limited availability of seeds from all accessions, quantification of amino acids, sugars and organic acids was performed for seeds of wild accession A2 and of modern bean accession M5. In general, the seed exudation profiles observed for modern accession M5 differed from those of wild accession A2 (Fig. 4). Regarding the exudation of amino acids, differences in decreasing order were observed for glutamate, aspartate, asparagine and alanine. For instance, the amount of glutamate in seed exudates of accession M5 was approximately 10 times higher than for accession A2. Also for several other amino acids, including glutamine, valine, threonine, among others, significant differences were observed between the wild and modern bean accessions (Figs. 4a and b). Similarly, the amount of fructose, glucose and sucrose was significantly higher for accession M5 than for accession A2 both at 24 and at 48h (Figs. 4c and d). No differences were observed in the amounts of organic acids (fumaric and malic acid) detected (Fig. 4e).

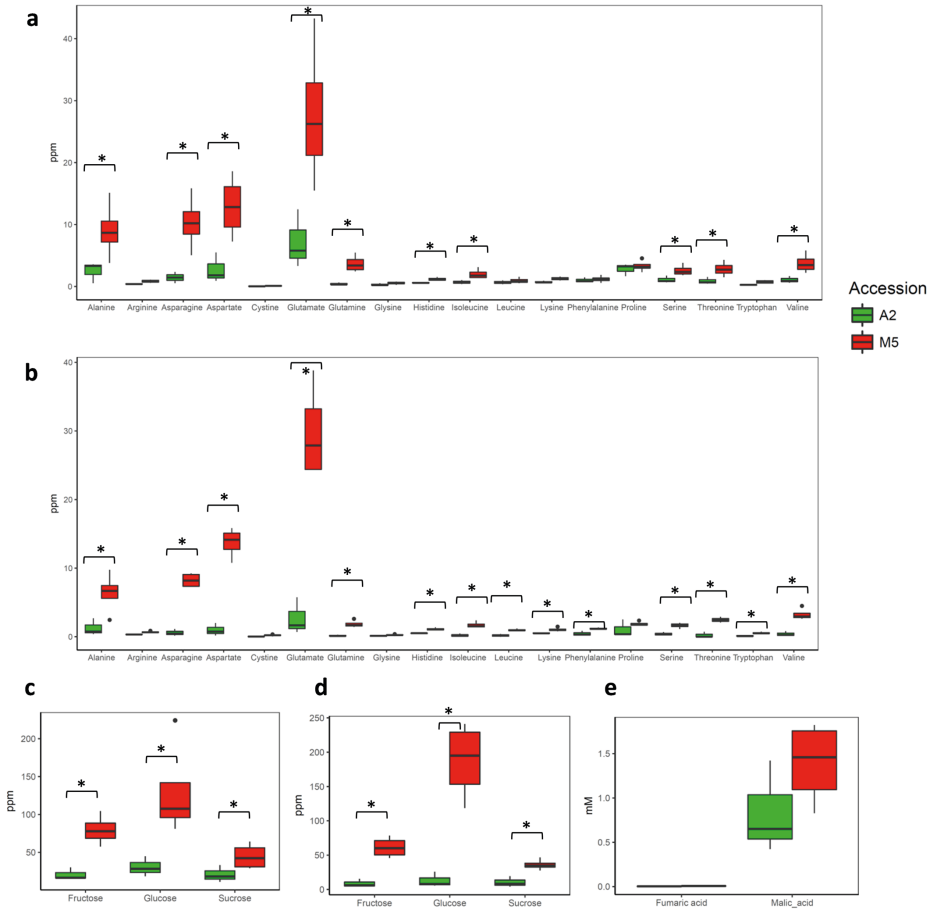


Fig. 4. Aminoacid, sugar and organic acid exudation by seeds of wild and modern common bean accessions A2 and M5. Box plots shows the amount of aminoacids (a and b) sugars (c and d) and organic acids (e), at 24h (a and c) and 48h (b, d and e) after water imbibition. Significant differences in the amount of exudates between the wild (A2 - green) and the modern (M5 - red) accessions were determined by the Wilcoxon signed-rank test ($P < 0,05$)

No clear effect of glutamate on bacterial community composition

Based on the observations that glutamate was the most abundant amino acid released 24 and 48h after seed imbibition, and that Actinobacterial taxa were significantly more abundant in the spermosphere of several modern bean accessions, we hypothesized that

glutamate enriches for Actinobacterial taxa. To begin to test this hypothesis, the same agricultural soil used for the spermosphere microbiome analyses was incubated successively in minimal media supplemented (or not) with glutamate. The results showed that soil samples treated with minimal media (M9) presented a very similar bacteria community composition (Fig. 5a) while soil samples treated with M9 supplemented with glucose, glucose + glutamate, and glucose + cysteine, presented high sample intra- and inter-variability (Fig. 5b). Therefore, the selective impact of glutamate on spermosphere microbiome composition was not apparent. Furthermore, it was observed that the addition of glucose, glutamate or cysteine led to a reduction of the relative abundance of the Actinobacteria as compared to the minimal medium and no specific and statistically significant effects of glutamate on the relative abundance of this phylum were found (Fig. S6).

The dominant bacterial phylum in all treatments was Proteobacteria with several abundant bacterial families dominating the community. In the samples enriched with glutamate, cysteine and glucose, most of the Proteobacterial sequences belonged to the *Pseudomonadaceae* family, while in the treatment with minimal media (M9) most of Proteobacterial sequences were assigned to the *Oxalobacteraceae* family (Figs. S6a and b). The family *Paenibacillaceae* (Firmicutes) was also found enriched as a consequence of the different treatments, especially in the treatment with cysteine. The family *Flavobacteriaceae* was found enriched in all the samples of the glucose treatment and also in a few samples from the glutamate and cysteine treatments; and finally, the family *Rhizobiaceae* was found enriched in almost all the samples from all the treatments; however, this enrichment was very pronounced in some samples, while for other samples within the same treatment the enrichment was minimal. In conclusion, the effect of glutamate on the relative abundance of members of the Actinobacteria phylum was not resolved with the experimental design used in this study.

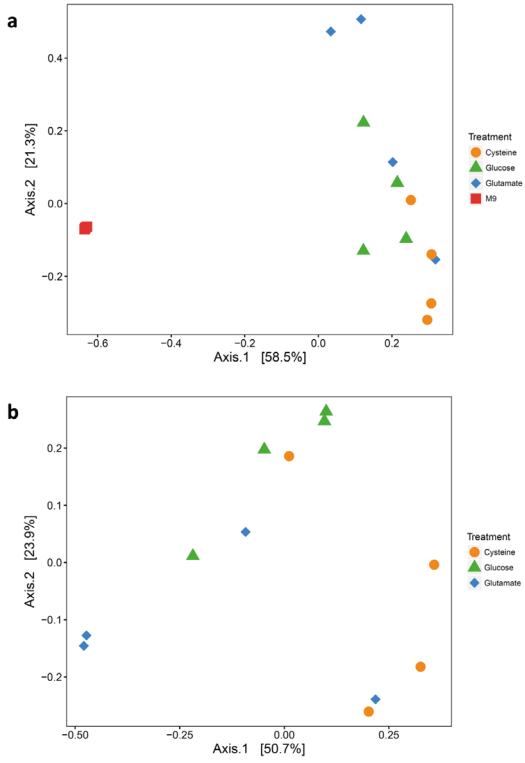


Fig. 5. Bacterial community structure of soil samples exposed to glutamate. To decipher the effect of glutamate on soil bacterial community assembly, four different treatments (n=4) were applied as follows: Soil enriched with M9 or minimal media (red boxes); soil enriched with cystein and glucose (yellow circles); soil enriched with glutamate and glucose (blue diamonds); and finally soil enriched with glucose alone (green triangles). Principal coordinate analysis of bray distances from MiSeq 16S sequence data were calculated, with (a) and without minimal media (M9) samples (b).

Table 1. Bacterial genera from the core isolated through cultivation methods. Taxonomic classification and the number of isolates obtained are shown.

Genus	Number of Isolates
Bacillus	41
Bradyrhizobium	8
Burkholderia-Paraburkholderia	18
Flavisolibacter	1
Lysobacter	33
Massilia	26
Mycobacterium	5
Paenarthrobacter	24
Pseudarthrobacter	17
Ramlibacter	7
Rhizobium	16
Sphingomonas	11
Streptomyces	66
Terrabacter	47

Discussion

In this study we characterized the spermosphere microbiome of wild and modern accessions of common bean grown in an agricultural soil from the Colombian highlands, a centre of common bean diversification. Firstly, we observed a spermosphere effect of seeds of wild and domesticated common bean accessions (*Phaseolus vulgaris*) on soil bacterial richness and evenness, represented by an overall lower bacterial diversity of the spermosphere compartment as compared to bulk soil. This reduction in bacterial diversity suggests that even during this short developmental stage, only 48 hours after sowing, a subset of soil bacterial taxa is enriched around the germinating bean seeds. This reduction in bacterial diversity in the spermosphere was also observed for plant species of the Brassicaceae family (Barret *et al*., 2015). Similarly, reduction of the bacterial richness has also been observed for the rhizosphere, the endosphere and the phyllosphere (Bulgarelli *et al*., 2015; Edwards *et al*., 2015; Ruiz-Pérez *et al*., 2016). The results of our study further showed that the domestication status has an impact, albeit subtle, on the

spermosphere microbiome composition of common bean. These results confirm and extend earlier observations we made for the rhizosphere microbiome of common bean (Pérez-Jaramillo *et al*. 2017). For wild and modern common bean accessions, the genera *Bacillus, Burkholderia-Paraburkholderia, Rhizobium, Massilia*, *Lysobacter, Paenarthrobacter, Luteimonas, Ramlibacter* and *Paenibacillus* were enriched as compared to bulk soil samples. The genera *Bacillus* and *Paenibacillus* have been described as colonizers of common bean seeds (López-López *et al*., 2010). Also the genus *Lysobacter* is common in agricultural soils (Puopolo *et al*. 2018) and known by its ability to produce a diverse set of secondary metabolites with antimicrobial properties that may aid in the control of soil-borne pathogens (Hayward *et al*., 2010; de Bruijn *et al*., 2015; Gómez-Expósito *et al*., 2015). Finally, the genera *Luteimonas* and *Ramlibacter* have also been found associated with different plant compartments (Pfeiffer *et al*., 2016; Wemheuer *et al*., 2017), however, their role as spermosphere abundant taxa remains to be elucidated. The most remarkable difference was the higher relative abundance of several Actinobacterial families in the spermosphere of modern bean accessions as compared to soil. At genus level, these differences in Actinobacterial taxa were mainly explained by a higher abundance of *Pseudarthrobacter*, *Paenarthrobacter* and *Streptomyces*. The genera *Pseudarthrobacter* and *Paenarthrobacter*, formerly classified within the genus *Arthrobacter*, are commonly found as natural soil inhabitants (Busse, 2016). Similarly, the genus *Streptomyces* is recognized as a conspicuous soil bacteria, many species within the genus are part of the plant microbiome and several strains have been reported as plant growth-promoting and effective biocontrol bacteria (Manteca and Sanchez, 2009; Viaene *et al*., 2016).

To what extent, seed endophytic bacteria (Berg and Raaijmakers, 2018) contributed to the genotype-dependent effect on the spermosphere bacterial microbiome composition is not yet known. The bacterial phyla Proteobacteria, Firmicutes and Actinobacteria have

been found as the main seed-associated bacteria for several plant species as well as diverse taxa within the fungal classes Dothideomycetes and Tremellomycetes (López-Velasco *et al*., 2013; Barret *et al*., 2015; Klaedtke *et al*., 2015; Rybakova *et al*., 2017). *Bacillus* and *Paenibacillus* have been also described as seed endophytic (López-López *et al*., 2010). Their high abundance in the spermosphere microbiome as compared to bulk soil may suggest that these seed endophytic bacteria rapidly colonized the spermosphere after imbibition, or that soil populations rapidly colonized the carbon rich spermosphere compartment. The first hypothesis is supported by the low abundances of both genera in bulk soil samples as well as by their very low abundances in rhizosphere samples (Pérez-Jaramillo *et al*., 2017). Most of these microbes are alive, and once the seed germinates, actively colonize the protruding seedling (López-Velasco *et al*., 2013; Adam *et al*., 2016). Under axenic conditions, the seed microbiome and the spermosphere microbiome might share a significant proportion of these endophytic microbial taxa. However, when the seed germinates in soil, the spermosphere microbiome is mainly composed by soil bacteria (Buyer *et al*., 1999). Although the impact of endophytic communities on spermosphere microbiome composition was not evaluated experimentally in our study, we postulate that the spermosphere compartment of common bean is colonized primarily by a subset of soil bacteria and that a small contribution of the endophytic microbiome may reinforce genotype-specific effects of the seeds on the spermosphere microbiome.

The pairwise comparison of the bacterial abundances in the spermosphere of wild and the modern common bean accessions showed that Actinobacterial families and genera were enriched in modern common bean. Despite that Actinobacteria is commonly found associated with plants, it is yet unclear why the association was more prominent with the spermosphere of modern bean. In a recent meta-analysis, we described that rhizospheres of wild plant species are enriched in Bacteroidetes, while Proteobacteria and Actinobacteria are enriched in the rhizosphere of modern plant cultivars (Pérez-Jaramillo

et al., 2018). As a possible explanation we previously proposed that both Actinobacteria and Proteobacteria may benefit more from the type and/or quantity of exudates that modern plants release (Pérez-Jaramillo *et al*., 2018). High competition for resources and space occurs within the spermosphere, and only microorganisms with traits for successful spermosphere competence may succeed. In this sense, fast-growing copiotrophic taxa possess a competitive advantage in the spermosphere compartment as compared to other resident soil bacteria (Lemanceau *et al*., 2017). In the bean spermosphere microbiome, a significant increase in members of the bacterial phyla Proteobacteria, Firmicutes and Bacteroidetes were observed. These groups have been associated with a copiotrophic life strategy (Fierer *et al*., 2007; Uksa *et al*., 2015; Ho *et al*., 2017). Conversely, Chloroflexi, Plactomycetes, Acidobacteria and Verrucomicrobia, generally regarded as oligotrophs (Ho *et al*., 2017), were significantly overrepresented in bulk soil samples. The most abundant genera in the core microbiome of the bean spermosphere belonged to the phylum Proteobacteria, represented by the genera *Sphingomonas*, *Burkholderia-Paraburkholderia, Bradyrhizobium, Massilia* and *Rhizobium*. The genus *Sphingomonas* has been reported as part of the core microbiome of the roots and phyllosphere of other plant species (Chen *et al*., 2018; Hamonts *et al*., 2018). The genus *Paraburkholderia*, also part of the core, is a recent genus that includes former non-pathogenic environmental *Burkholderia* (Eberl and Vandamme, 2016). This genus has been found associated with legume plants as nitrogen fixing bacteria and also was found as the predominant genus within the nodules of common bean plants grown in natural soils (Dall'Agnol *et al*., 2016). The occurrence of *Paraburkholderia* in the core spermosphere microbiome of common bean, together with *Bradyrhizobium* and *Rhizobium*, which are bacterial genera commonly associated with common bean nodules (Aguilar *et al*., 2004; Wang *et al*., 2016), suggests that bacterial genera with nitrogen fixing capabilities are actively recruited to the bean spermosphere compartment, irrespective of the domestication status.

Finally, the genus *Massilia* has been found colonizing seeds and roots of several plant species, and several beneficial traits related with biological control have been described, such as the production of siderophores and extracellular lytic enzymes (Ofek *et al*., 2012). In synthesis, we postulate that the main taxa in the spermosphere core microbiome of common bean are copiotrophic bacteria that may use seed exudates as a carbon source, can establish symbiotic associations with the plant and have the potential to inhibit soil-borne plant pathogens.

In the original definition by Onorato Verona (1958), the spermosphere is regarded as a zone of intense microbial activity around a germinating seed, as a result of carbon deposition (Nelson, 2004). Therefore, the spermosphere microbiome assembly relies, at least in part, on the composition of the seed exudates released by seeds during germination (Lemanceaeu *et al*., 2017). For roots, differences on exudation profiles between wild and modern plant genotypes have been already reported. For instance, wild and primitive genotypes of wheat showed differences in root exudation as compared to modern wheat varieties. Accordingly, higher contents of fructose, galactose and *myo*-inositol were observed in wild and primitive genotypes, while higher contents of isomaltose, sucrose, hexadecanoic acid, octadecanoic acid and 1-octaconasol were detected in modern wheat varieties (Iannucci *et al*., 2017). In this study we observed that common bean seeds released several amino acids, organic acids and sugars during imbibition and germination. The main difference in seed exudation profiles between wild and modern accessions was related with the quantity of the compounds released by the seed. We observed higher quantities of various exudate constituents by the modern accession M5 as compared to wild accession A2. It is important to emphasize that the collection of seed exudates was performed *in vitro*, while the characterization of the spermosphere microbiome was performed directly in soil. The exudate composition observed for the two accessions after 48 hours may not directly correspond to the exudate composition in the soil after the same

time, and in consequence a particular spermosphere microbiome composition may not be the consequence of a particular set of exudates detected in vitro. This may have affected the results of the assay in which soil was treated with different amino acids minimal media and glucose as a carbon source. As glutamate was the most abundant amino acid released by modern bean accessions *in vitro*, and Actinobacterial taxa were significantly more abundant in the spermosphere microbiome, we hypothesized that glutamate could have a positive effect on the abundance of actinobacterial taxa. What we observed was that the Phylum Actinobacteria responded negatively to the treatments, while Proteobacteria and to a lesser extent Firmicutes, responded positively. Most probably, the concentration of glucose was optimal for the growth of competitive and fast growing taxa within Proteobacteria and Firmicutes. In order to accurately evaluate the effect of glutamate on soil bacterial communities, an optimization of glucose and amino acid levels in the culture media is needed. Furthermore, it is pivotal to validate whether glutamate is also a more abundant seed exudate in the spermosphere of modern accessions germinating in soil.

Finally, 66.6% of the genera identified as part of the core spermosphere microbiome were succesfully isolated by common culture-dependent methods. Furthermore, it was possible to match MiSeq V3-V4 metagenome sequence data with the V1-V4 sequence data from the isolate collection (Table S1). This isolate repository will be tested in future experiments, either alone or in consortia, for their ability to colonize the spermosphere of common bean accessions and to utilize specific exudate components as well as for their effects on growth and health of common bean.

Conclusions

The spermosphere is an active and dynamic compartment that, despite its short duration, can impact the diversity and abundance of soil bacterial communities. A subtle but significant effect of the domestication of common bean was revealed in the spermosphere

microbiome composition, where modern accessions showed a higher abundance of several Actinobacterial taxa. The spermosphere core microbiome was composed of several copiotrophic bacteria, with several bacteria capable of fixing nitrogen. Additionally, several taxa within the core spermosphere microbiome have been reported as plant growth promoting bacteria and with biocontrol activities. Several bacterial genera found in the core were isolated through culture-dependent methods and the resulting bacterial repository will be the basis for future experiments to validate their differential ability to colonize the spermosphere of wild and modern bean accessions. Finally, it was not possible to validate *in vitro* the role of glutamate, which was the most abundant amino acid detected in seed exudates of the modern bean accession M2. Optimization of the culture media and glucose levels is needed in order to accurately determine the selective impact of glutamate on soil bacterial communities, especially for Actinobacterial taxa.

Supplementary materials

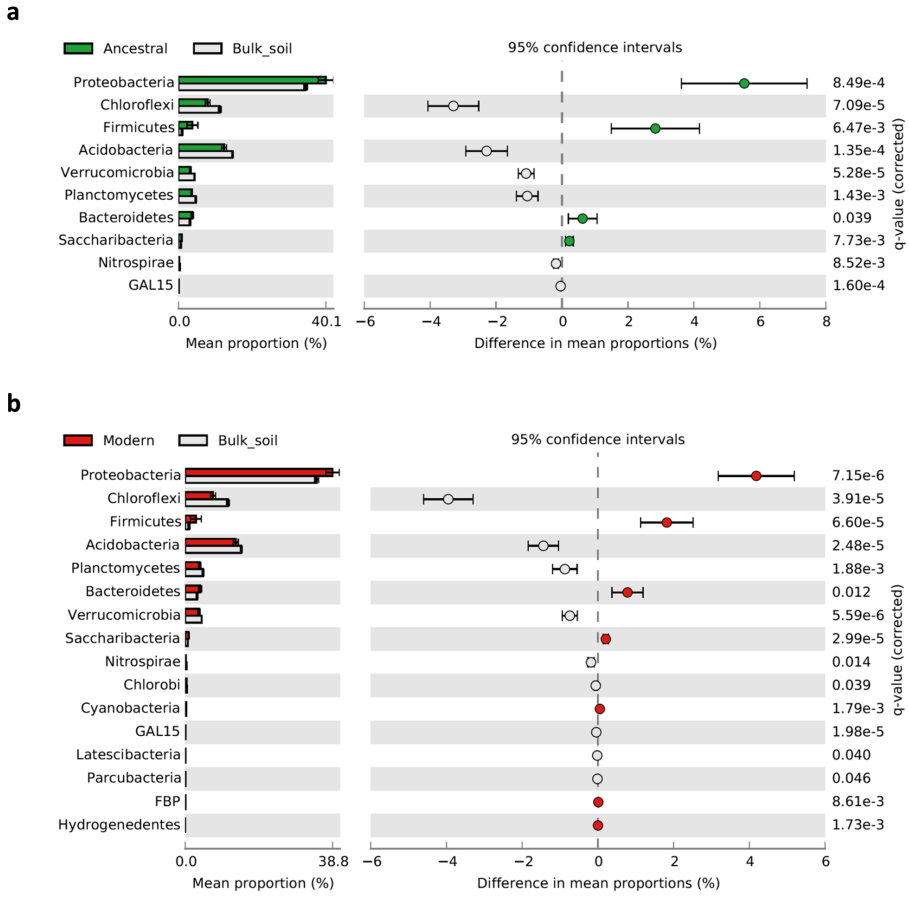


Fig. S1. Differential abundance of bacterial phyla in spermosphere and bulk soil samples. Extended error bar plots showing differentially abundant phyla between (a) the spermosphere microbiome of ancestral and (b) modern common bean accessions and bulk soil samples. Welch's t-tests followed by BH corrections were performed between merged spermosphere samples by domestication status and bulk soil samples. Only differentially abundant Phyla are shown.

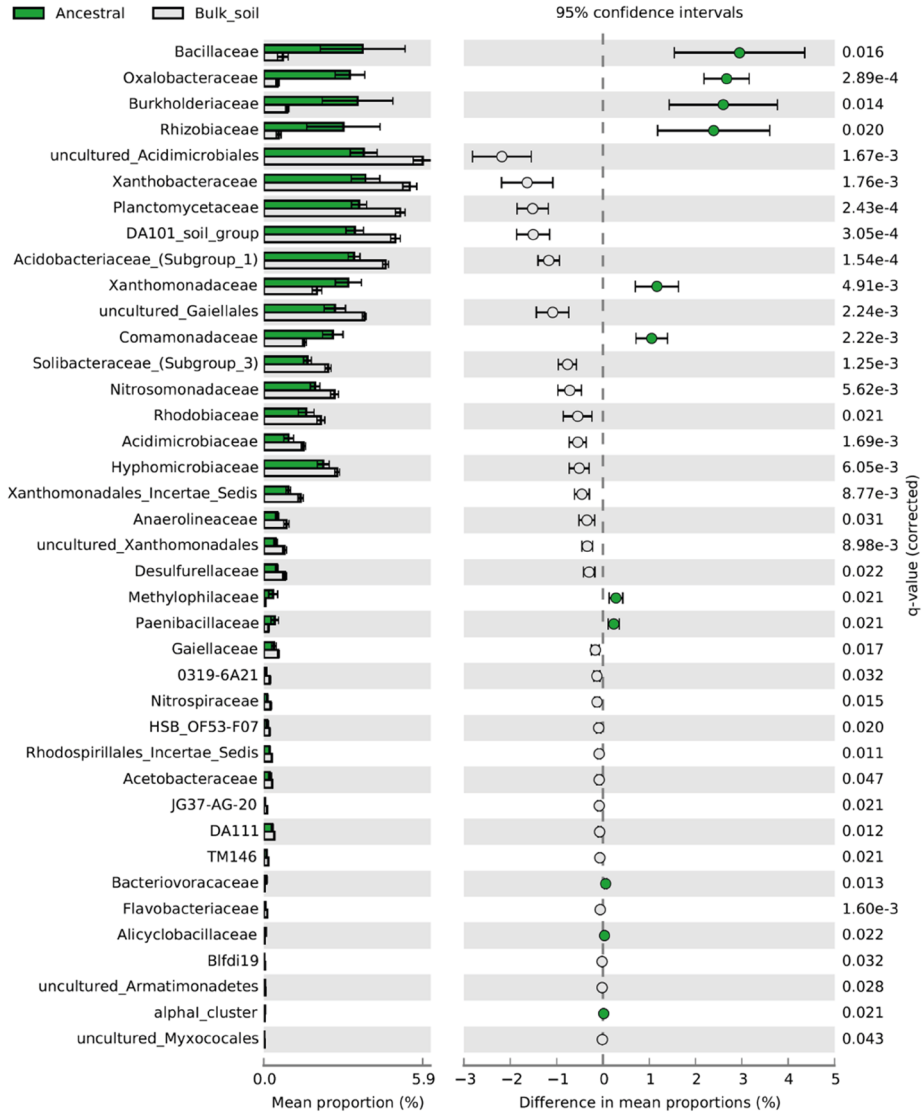


Fig. S2. Differential abundance of bacterial families in the spermosphere of ancestral common bean and bulk soil samples. Extended error bar plots showing differentially abundant families between the spermosphere microbiome of ancestral common bean accessions and bulk soil samples. Welch's t-tests followed by BH corrections were performed. Only differentially abundant families are shown.

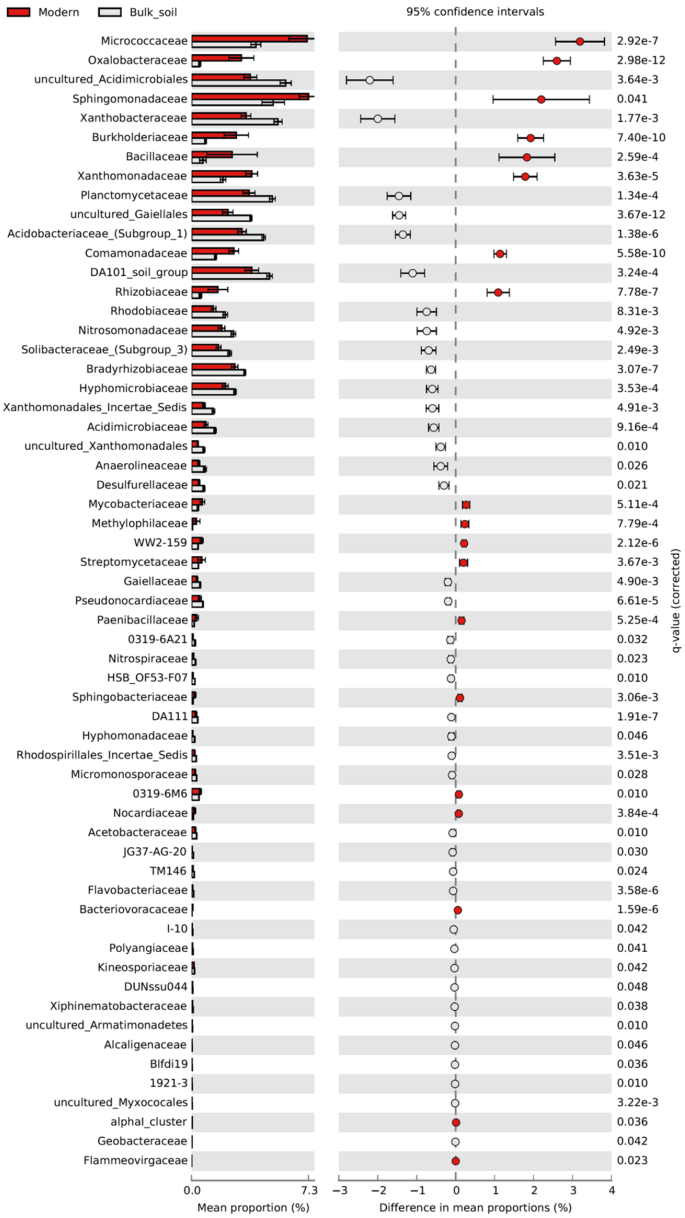


Fig. S3. Differential abundance of bacterial families in the spermosphere of modern common bean and bulk soil samples. Extended error bar plots showing differentially abundant families between the spermosphere microbiome of modern common bean accessions and bulk soil samples. Welch's t-tests followed by BH corrections were performed. Only differentially abundant families are shown.

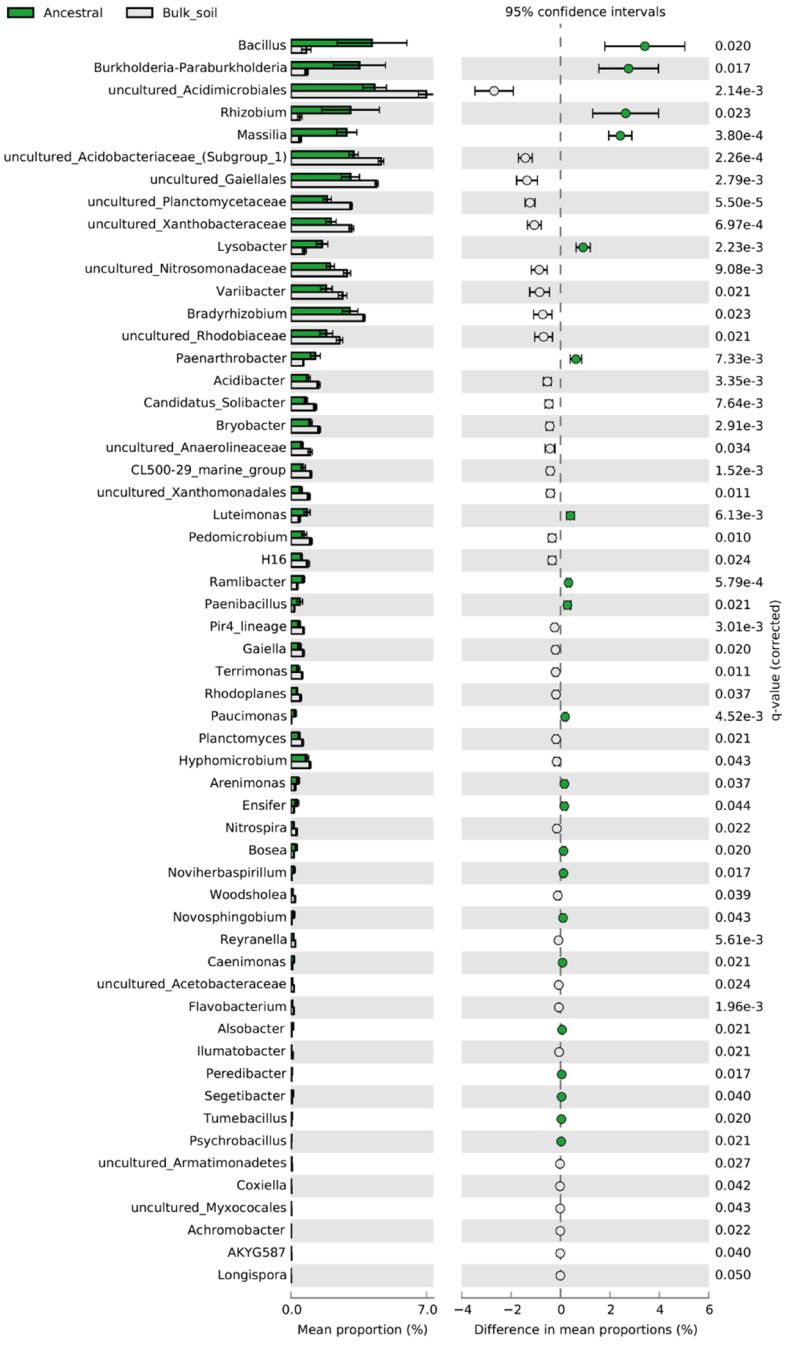


Fig. S4. Differential abundance of bacterial genera in the spermosphere of ancestral common bean and bulk soil samples. Extended error bar plots showing several differentially abundant genera between the spermosphere microbiome of ancestral common bean accessions and bulk soil samples. Welch's t-tests followed by BH corrections were performed.

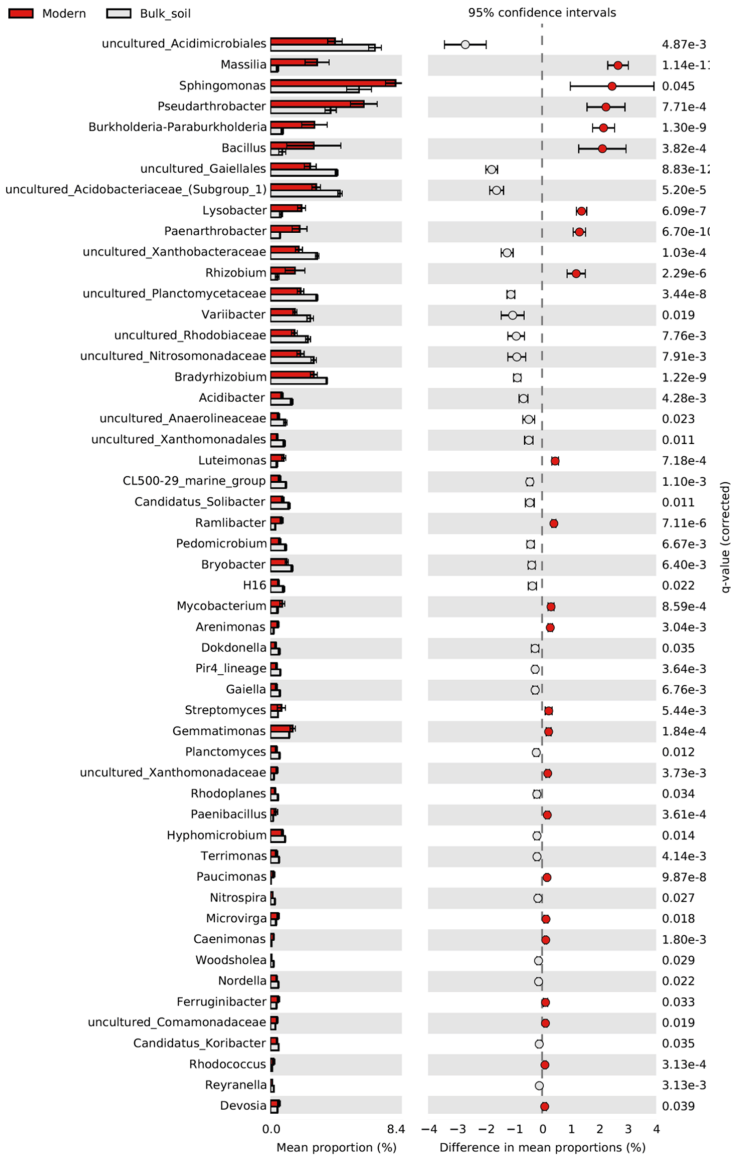


Fig. S5. Differential abundance of bacterial genera in the spermosphere of modern common bean and bulk soil samples. Extended error bar plots showing several differentially abundant genera between the spermosphere microbiome of modern common bean accessions and bulk soil samples. Welch's t-tests followed by BH corrections were performed.

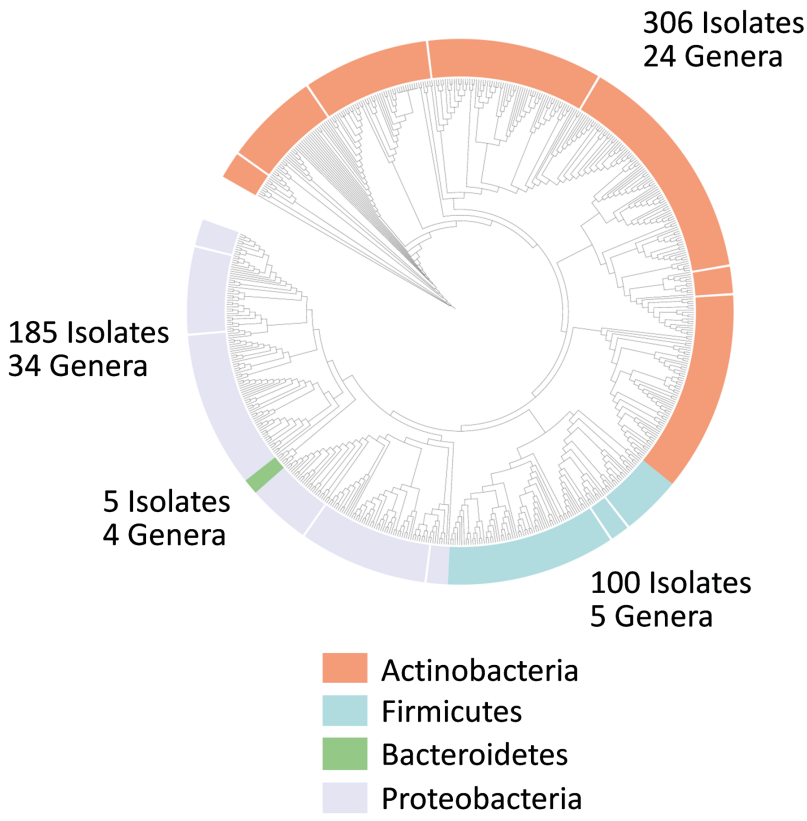


Fig. S6. Phylogenetic tree of bacterial strains isolated from common bean spermosphere. The four bacterial phyla identified are shown with colors within the tree. Additionally, the number of genera and isolates per bacterial phyla are indicated.

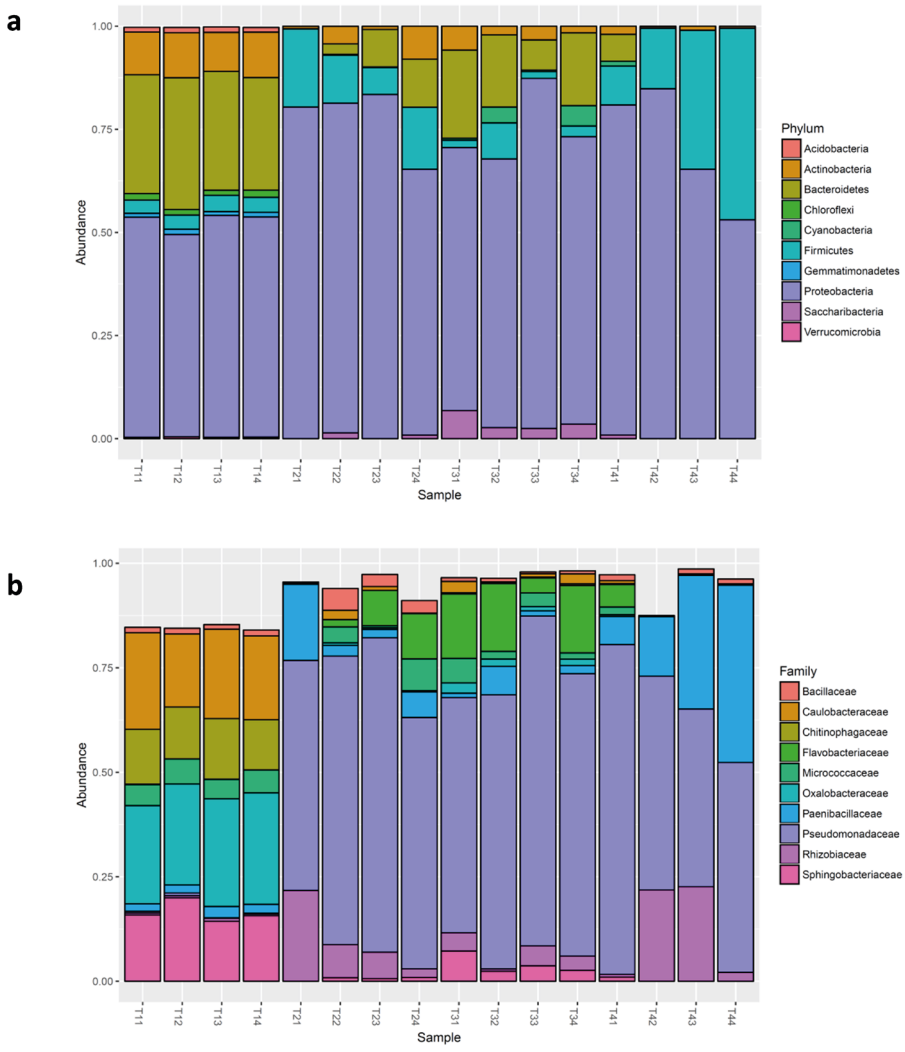


Fig. S7. Bar graphs of the most abundant phylum and families in soil under treatment with different amino acids minimal media and glucose. T1 (minimal media treatment) consisted of M9 minimal media as a control treatment; T2 (glutamate treatment) consisted of M9 minimal media, 10 mM of glutamate, 1 mM of $MgSO_4$, and glucose 20%; T3 (glucose treatment) consisted of M9 minimal media, 1 mM of $MgSO_4$, and glucose 20%; and T4 (cysteine treatment) consisted of M9 minimal media, 10 mM of cysteine, 1 mM of $MgSO_4$, and glucose 20%. The numbers following the treatment correspond to the replicate number.

Table S1. Match of MiSeq V3-V4 spermosphere data and the V1-V4 Sequences from the isolate collection. Taxonomic classification, the number of reads obtained with MiSeq sequencing, and the number of isolates retrieved per OTU_ID is shown.

Phylum	Genus	OTU_ID	Spermosphere MiSeq reads	Isolates matched
Proteobacteria	*Sphingomonas*	OTU_2	67739	1
Actinobacteria	*Pseudarthrobacter*	OTU_4	61652	2
Proteobacteria	*Bradyrhizobium*	OTU_3	27957	1
Gemmatimonadetes	uncultured_*Gemmatimonadaceae*	OTU_18	20328	**
Actinobacteria	*Paenarthrobacter*	OTU_14	19375	1
Proteobacteria	uncultured_*Xanthobacteraceae*	OTU_12	19164	**
Firmicutes	*Bacillus*	OTU_42	16161	2
Proteobacteria	uncultured_*Rhodobiaceae*	OTU_22	13817	**
Actinobacteria	*Terrabacter*	OTU_30	12070	2
Acidobacteria	*RB41*	OTU_37	11979	**
Proteobacteria	*Rhizobium*	OTU_9	8647	1
Proteobacteria	unclassified_*JG34-KF-161*	OTU_6***	6819	2
Proteobacteria	*Devosia*	OTU_20	5333	1
Actinobacteria	*Streptomyces*	OTU_8	4660	1
Proteobacteria	*Rhizobium*	OTU_1	3528	2
Proteobacteria	*Rhizobium*	OTU_253	2004	2
Proteobacteria	unclassified_*Comamonadaceae*	OTU_11	1057	**
Proteobacteria	unclassified_*Comamonadaceae*	OTU_21	972	1
Bacteroidetes	*Dyadobacter*	OTU_5*	826	1
Proteobacteria	unclassified_*Sphingomonadaceae*	OTU_16****	740	1
Bacteroidetes	*Terrimonas*	OTU_1424,OTU_1906,OTU_2003,OTU_202,OTU_979	639	1

*With MiSeq/Sanger matched with several OTUs: OTU_10611,OTU_11077 and OTU_5.
**Not available in the collection
***With MiSeq/Sanger as: Sphingomonadaceae; *Sphingomonas*.
****With MiSeq/Sanger as: Sphingomonadaceae; *Novosphingobium*.

Chapter 7

General discussion

The interactions between plants and microorganisms and the consequences of these interactions for plant growth and health have promoted a field of intense research. Currently, it is known that the plant-associated microbiome can provide a plethora of beneficial services, such as disease suppression, induction of systemic resistance, enhanced nutrient acquisition and increased tolerance to biotic and abiotic stresses (Mendes *et al*., 2011; Berendsen *et al*., 2018; Carrión *et al*., 2018; Hassani *et al*., 2018). Several root physiological characteristics and spatio-temporal differences in root exudation profiles have been shown to structure, at least in part, the rhizosphere microbiome (Jones *et al*., 2009; Sasse *et al*., 2017). To date, however, little is known on if and how plant domestication and the loss of genetic diversity affected the ability of plants to successfully assemble a beneficial microbiome in the rhizosphere and endosphere. The overall aims of my thesis were to i) assess the effect of domestication of common bean (*Phaseolus vulgaris*) on rhizosphere and spermosphere microbiome diversity and structure; ii) link specific root architectural traits and exudation profiles to the abundance of specific bacterial groups, and iii) evaluate how habitat expansion impacted rhizosphere microbiome assembly of common bean. In this chapter, I discuss the major results obtained as well as the different future research avenues.

Revisiting the impact of plant domestication on microbiome composition

It is well stablished that plant domestication and subsequent improvement of crop cultivars by genetic breeding caused phenotypic, genomic and metabolic changes which enabled humans to use several plant species for consumption and large-scale food production (Flint-García, 2013; Abbo *et al*., 2014; Smýkal *et al*., 2018). Many of these changes were accompanied by other inadvertent effects, such as the reduction in the genetic diversity of domesticated crop cultivars (Doebley *et al*., 2006) and the reduced ability of the domesticated crops to cope with herbivorous insects (Chen *et al*., 2015).

Regarding the impact of plant domestication on symbiotic associations between plants and soil microbes, it has been shown that mycorrhizal and rhizobial interactions were, in some cases, undermined (Hetrick *et al*., 1992; Kiers *et al*., 2007; Lehmann *et al*., 2012; Sangabriel-Conde *et al*., 2015; Ellouze *et al*., 2018; **Chapter 2**). To date, however, the effect of plant domestication on the vast majority of microbes that live in the spermosphere, rhizosphere, phyllosphere or endosphere is largely unknown. The first attempts to unveil the impact of domestication on root and rhizosphere microbial communities were performed through culture-dependent methods. For instance, Germida and Siciliano (2001) showed that the rhizosphere bacterial community of ancient landraces of wheat was more diverse than that of two modern wheat cultivars. Similarly, characterization of the endophytic bacteria in stems and roots of wild and cultivated rice showed that the number of diazotrophic endophytes was higher in wild rice than in cultivated rice (Engelhard *et al*, 2000; Koomnok *et al*, 2007). Subsequent profiling of rhizospheric bacterial and fungal communities of the wild relative of maize, *Zea mays* subsp. *parviglumis* (Balsas teosinte) and two domesticated maize cultivars showed that several bacterial diversity indices were higher for teosinte than for the domesticated cultivars (Szoboszlay *et al*., 2015). Recently, with the advent of high throughput (HTP) sequencing technologies, the taxonomic diversity of plant-associated microbiomes of several wild and modern cultivars have been determined. Most of these studies showed that the α-diversity of the plant-associated microbiomes was similar for domesticated and wild plant species (Bulgarelli *et al*., 2015; Cardinale *et al*., 2015; Shenton *et al*., 2016; Chaluvadi and Bennetzen, 2018). Nevertheless, in a study with *Agave* species, cultivated species showed a lower prokaryotic diversity than the two wild species (Coleman-Derr *et al*., 2016). Also in a study with rice, wild species showed the highest diversity of root-associated bacteria as compared to modern rice (Tian *et al*., 2017). Remarkably, it has been repeatedly found that the structure of the microbial communities associated with

wild and cultivated plant species diverges as well as the abundance of certain microbial groups, which responded differently depending on the domestication status of the plant: in many cases, the abundance of certain taxa was reduced or augmented in domesticated/wild plants. Studies conducted with sugar beet (Zachow *et al*., 2014), barley (Bulgarelli *et al*., 2015) and lettuce (Cardinale *et al*., 2015) showed an enrichment of members of the Bacteroidetes in the rhizosphere of the wild relatives as compared to their domesticated counterparts (**Chapter 5**). Similarly, the bacterial genera *Flavobacterium* and *Pedobacter*, both from the Bacteroidetes phylum, were enriched in wild rice as compared to cultivated rice (Tian *et al*., 2017). For the rhizosphere of wild rice, the study by Shenton *et al*. (2016) further revealed an enrichment of bacterial taxa of the Anaerolineae and Nitrospirae families and a depletion of Saprospirae taxa as compared to rhizosphere of domesticated rice (Shenton *et al*., 2016). A differential abundance of root and rhizosphere bacterial taxa was also observed for foxtail millet (*Setaria italica*) and its wild ancestor green foxtail (*S. viridis*): while Gammaproteobacteria, Deltaproteobacteria and Firmicutes were enriched in the rhizosphere of the wild ancestor, Alphaproteobacteria was enriched in the rhizosphere of foxtail millet (Chaluvadi and Bennetzen, 2018). In contrast, Leff *et al*. (2016) did not observe a significant effect of the sunflower genotype on the assembly of rhizosphere bacterial communities but did find effects on fungal rhizosphere community composition. Accordingly, it was observed that unclassified *Chaetomiaceae*, *Olpidium* and *Mortierella spp.* had higher relative abundances in wild sunflower while modern sunflower accessions had higher relative abundaces of *Pleosporales*, *Preussia* spp., unclassified *Thelebolaceae*, *Fusarium* spp., and *Conocybe* spp. (Leff *et al*., 2016). In summary, there is an increasing number of studies that have shown a significant impact of plant domestication on the rhizosphere microbiome composition. To what extent plant phenotypic, genotypic and physiological

alterations caused by domestication have contributed to these microbiome changes is yet unknown.

Phaseolus vulgaris domestication and its impact on microbiome assembly

In this thesis common bean was used as a plant 'model' to assess the impact of domestication on the diversity of root- and seed-associated bacterial communities. Common bean (*Phaseolus vulgaris* L.) is a very important legume crop, grown worldwide in a broad range of environmental conditions (Broughton *et al*., 2003). Common bean is a good source of protein, carbohydrate and micronutrients and represents the basis of subsistence economy for low income farmers in developing countries (Akibode and Maredia, 2011; CGIAR, 2013). Initially, using an agricultural soil collected in the Colombian highlands, I assessed the rhizosphere bacterial assembly and characterized root phenotypic traits of wild and modern common bean accessions. In summary, we observed a higher relative abundance of Bacteroidetes, mainly *Chitinophagaceae* and *Cytophagaceae*, in the rhizosphere of wild bean accessions and an increase in relative abundance of Actinobacteria and Proteobacteria in the rhizosphere of modern bean accessions. It was also possible to associate the divergences in rhizobacterial community composition between wild and modern bean accessions with differences in root morphological traits. Wild common bean accessions showed significantly higher specific root length (SRL) values as compared to modern bean accessions. Subsequently, it was shown that the SRL explained a significant portion of the variability in rhizosphere microbiome observed, specifically the abundance of families from the phylum Bacteroidetes (**Chapter 3**). Therefore we concluded that the higher relative abundance of the families *Chitinophagaceae* and *Cytophagaceae* in wild common bean accessions is explained, at least in part, by the higher SRL displayed by these accessions. It has been postulated that a high SRL may provide a higher efficiency of water search and uptake,

traits that in wild beans may be pivotal to prosper and survive in dry native habitats (Toro, 1990; Comas *et al*., 2013). The number of nodules also explained a small portion of the total variability, and it was specifically related with subtle dissimilarities between the two wild bean accessions A1 and A2. For the legume plant *Lotus japonicus,* key symbiotic genes were shown to play a major role in the assembly of bacterial communities in root and rhizosphere compartments (Zgadzaj *et al*., 2016). Whether the symbiotic plant-microbe associations are also structuring forces of the rhizobacterial community structure in common bean remains to be investigated.

Using the same computational pipeline for processing 16S raw data from several studies, higher relative abundances of members of the phylum Bacteroidetes were also found for wild relatives of several other crop species, whereas the rhizosphere of domesticated plant accessions was more enriched in members of the Actinobacteria and Proteobacteria phyla (**Chapter 5**). Why Bacteroidetes are relatively more abundant in the root and rhizosphere compartments of wild relatives of various plant species is yet unknown. We hypothesized that their prevalence in the root compartments of wild plant species may be a phylogenetic signal associated with the presence of complex biopolymers in their root exudates. Differences in plant root exudates and root architecture between crop plants and their wild relatives may be the pivotal drivers of rhizospheric interactions (Saleem *et al*., 2018; Sasse *et al*., 2018). However, very little is known about the differences in exudation profiles between wild and modern plant genotypes. For wheat, it was shown that a modern variety exuded three to five times more “simple” sugars (mainly fructose, glucose, and maltose) than an ancient wheat cultivar under stress conditions (Shaposhnikov *et al*., 2016). Also for wheat higher contents of fructose, galactose and *myo*-inositol were observed in wild and primitive genotypes, while higher contents of isomaltose, sucrose, hexadecanoic acid, octadecanoic acid and 1-octaconasol were detected in modern wheat

varieties (Iannucci *et al*., 2017). Similarly, it is known that several wild plant species present differences in root architectural traits, such as root density and SRL, as compared to modern cultivars (Martín-Robles *et al*., 2015). However, very little is yet known about differences in seed and root exudation profiles as well as root morphological traits between wild relatives and modern cultivars of other plant species. Recent metabolomics techniques and analyses can be used to accurately determine the composition of exudates released by wild relatives and their modern counterparts, including targeted and nontargeted approaches. In this way, both known and unknown molecules can be captured to expand the catalogue of compounds released by plant seeds and roots.

Domestication not only encompasses changes in plant genetics and physiology but also involves habitat expansion and agricultural management. To begin to understand how these domestication factors affect the rhizosphere microbiome, the bean accessions were grown in a native and an agricultural soil from the Colombian highlands (**Chapter 4**). In general, the results indicated that the transition of common bean from a native soil to an agricultural soil led to a gain of rhizobacterial diversity and led to a stronger bean genotype-dependent effect on rhizosphere microbiome assembly. A core rhizosphere microbiome was identified and mainly consisted of genera from the Proteobacteria phylum, in particular *Rhizobium*, *Bradyrhizobium, Mesorhizobium, Burkholderia, Novosphingobium* and *Sphingomonas*. A significant portion of this common bean core microbiome is composed of bacterial genera with nitrogen fixing capabilities, an important feature of microbes associated with leguminous plant species. For other non-leguminous plant species, these bacterial genera are also members of the core rhizosphere microbiome (Peiffer *et al*., 2013; Yeoh *et al*., 2016; 2017), suggesting a homogenization of rhizobacterial diversity of plants grown in different agricultural landscapes. Co-occurrence analyses further showed a reduction in complexity of the interactions going

from native to agricultural soil. Based on these results I concluded that rhizobacterial community assembly for common bean grown in agricultural soil is less complex and more modular than for common bean in native soil, making it relatively more easy for a given soil bacterial species to invade and establish in the rhizosphere of bean plants grown in agricultural soils.

For the spermosphere, we observed a decrease of α-diversity for all bean accessions as compared to bulk soil, which resembles the reduction in diversity revealed for other plant compartments (Bulgarelli *et al*., 2015). Proteobacteria, Actinobacteria and Firmicutes were enriched in the bean spermosphere and small but significant differences in β-diversity were detected between the spermosphere microbiomes of wild and modern accessions. These results suggested a domestication effect on microbiome assembly already at this early developmental stage of common bean. Exudation profiles showed that the spermosphere of a modern bean accession contained higher levels of glutamate and glutamine as compared to a wild bean accession. Therefore an *in vitro* assay was performed to evaluate whether glutamate exerts a selective effect on the abundance of families from the Actinobacteria phylum. Although a strong impact on microbiome composition was observed, it was not possible to observe a clear trend towards specific bacterial taxa consistently enriched by the amino acid treatment. More specifically, the addition of glucose, glutamate or cysteine led to a reduction of the relative abundance of the Actinobacteria as compared to the minimal medium. Statistically significant effects of glutamate on the relative abundance of this phylum were not found. In order to accurately evaluate the effect of glutamate on soil bacterial communities, an optimization of glucose and amino acid levels in the culture media is needed. Furthermore, it is pivotal to validate whether glutamate is also a more abundant seed exudate in the spermosphere of modern accessions germinating in soil.

Common ground and next steps in plant microbiome research

In this thesis I aimed to decipher the impact of plant domestication on the spermosphere and rhizosphere microbiome composition of common bean. In an effort to go a step further I tried to link those particular assemblies with plant genotypic and phenotypic traits, in particular root phenotypic traits and specific exudates of wild and modern accessions of common bean (**Chapters 3 and 6**). Several exudates and hormones such as malic acid (Lakshmanan *et al*., 2012), citric acid (Zhang *et al*., 2014), salicylic acid (Lebeis *et al*., 2015; Berendsen *et al*., 2018), ferulic acid (Beckers *et al*., 2016), coumarins (Stringlis *et al*., 2018) as well as plant emitted volatile organic compounds (Rasmann *et al*., 2005; Schulz-Bohm *et al*., 2018) have been proposed as modulators of plant microbiome assembly. In a recent study, root exudates collected from *Avena barbata* were used to grow *in vitro* several typically rhizosphere-associated bacterial isolates and the authors observed a preference of these bacteria for consumption of aromatic organic acids such as nicotinic, shikimic, salicylic, cinnamic and indole-3-acetic (Zhalnina *et al*., 2018). In this thesis, it was observed that modern accessions of common bean showed higher exudation of the amino acid glutamate and I tried to link the higher abundance of this amino acid with the abundance of Actinobacterial families (**Chapter 6**). Despite these efforts, and that several exudates have been proposed as pivotal microbial regulators, validation that the exudate is produced and released into the soil by plant roots is required. Most exudate characterizations are performed in hydroponic or *in vitro* systems, without taking into account the possible impact of soil structure and texture on exudation profiles. In soil, root exudates are rapidly catabolized by microorganisms in the rhizoplane and the rhizosphere. Therefore, the first step would be to use specific devices such as microsuction cups to sample root exudates directly from plant roots growing in soil (Dessureault-Rompré *et al*., 2006; Eisenhauer *et al*., 2017). Subsequently, it is important to use both targeted and untargeted metabolite analyses (Bingol, 2018) in order to

characterize the broad range of root exudates, beyond the already known sugars, organic acids and amino acids that are commonly determined.

Similarly, the influence of root architecture and particular root physiological traits on microbiome assembly is still largely unknown. In this thesis, I linked the specific root length of wild common bean accessions with the abundance of the bacterial families *Chitinophagaceae* and *Cytophagaceae* from the Bacteroidetes Phylum. However, the mechanism(s) underlying the significant association of these Bacteroidetes families with roots of wild common bean was beyond the reach of the experimental setup used in this thesis. Differences in root morphological traits were also described for maize, and the root system of teosinte, which is the maize ancestor, had proportionally more fine (diameter < 0.03 mm) roots than modern varieties and it developed the highest root to shoot dry weight ratio (Szoboszlay *et al*., 2015). Thinner roots may provide a higher efficiency of water search and uptake, characteristics that are of pivotal importance in dry periods (Comas *et al*., 2013). How these root phenotypic traits affect the recruitment and colonization of microbes is largely unknown (Saleem *et al*., 2018). Additionally, most root phenotyping methods are based on seedling plants, and the extrapolation of this data to predict the architecture of a mature root system is problematic (Zhao *et al*., 2017). Therefore, the first step to decipher the effect of root architectural traits on microbiome assembly is to perform phenotyping of root systems in rhizotrons suitable for plant growth in controlled conditions and non-invasive image acquisition of root systems (RhizoTubes) (Jeudy *et al*., 2016). Subsequently, and using the same phenotyping system, sampling of specific root regions, such as root tips, lateral roots and mature root zones, should be performed to get a better insight into the spatial distribution of root microbiomes. Characterization of microbial communities through metataxonomic and metagenomic approaches associated with these regions, in plants grown expressing their natural root

plasticity, can unveil the diversity and functional repertoire of microbes as a function of root phenotypic traits.

The core microbiome and the search for heritable traits

In this thesis we also aimed to find bacterial taxa differentially abundant in wild common bean and in modern accessions. Additionally, we looked for a particular set of microbes that could be part of a microbial core. Accordingly, we found that the bacterial genera *Rhizobium*, *Bradyrhizobium, Mesorhizobium, Sphingomonas,* and *Streptomyces* were part of highly abundant taxonomic core that was shared by all eight bean accessions in the agricultural and in the native soils tested in this thesis. A significant portion of the core was composed of bacterial genera with nitrogen fixing capabilities. Several studies have also characterized the core microbiome for other non-leguminous plant species and the same rhizobacterial genera that made part of the common bean core were found as abundant core members (Peiffer *et al*., 2013; Yeoh *et al*., 2016; 2017; Chen *et al*., 2018; Walters *et al*., 2018). Recently, it has been proposed that core microorganisms may help organizing the assembly of resident plant-associated microbiomes within and around host plants (Toju *et al*., 2018). In the same perspective paper, it was proposed that one strategy for plant microbiome management is the core reinforcement strategy. Briefly, it is proposed that co-inoculations of core microbes with complementary roles may help increase plant physiological homeostasis, promoting stable establishment of late colonizing microorganisms (Toju *et al*., 2018). In this sense, future research needs to focus on the isolation and functional characterization of microbial species that are part of the core, and on the evaluation of how the re-introduction of these core microbiome members helps plants to overcome biotic and abiotic stresses. For common bean, I managed to establish an initial bacterial collection that was highly representative for the dominant taxa identified in the core spermosphere. These results suggest that a substantial

proportion (66.6%) of the bacteria detected as part of the core microbiome by NGS, presenting relative abundances higher than 0.5%, are indeed culturable contradicting the old paradigm that only 1-5% of the microbes detected in soil and plant-associated environments is culturable.

Finally, the search for heritable plant traits that can be exploited in breeding strategies is one of the main priorities in present and future plant microbiome research. One of the first attempts to identify plant traits and genes involved in beneficial interactions with soil microorganisms was done using recombinant inbred line (RIL) populations, derived from a cross between a cultivated tomato and a wild tomato relative (Smith *et al*., 1999). In that study, it was shown that three quantitative trait loci (QTL) were associated with disease suppression by the biocontrol strain *Bacillus cereus* UW85, and three other QTL were associated with the growth of *B. cereus* on the seed (Smith *et al*., 1999). A more recent study provided, for the first time, insights into genomic regions in the model plant Arabidopsis associated with phyllosphere microbiome composition. More specifically, loci responsible for defense and cell wall integrity were linked to variations in the composition of the microbiome of Arabidopsis leaves (Horton *et al*., 2014). Furthermore, genetic loci have been identified for their involvement in the ability of Arabidopsis accessions to respond to the activity of plant growth-promoting bacteria (Wintermans *et al*., 2016). Mendes *et al*. (2018) showed that resistance breeding in common bean against *Fusarium oxysporum* also had unintentionally co-selected for plant traits that support rhizosphere microbes with potential antagonistic activities to this fungal root pathogen. However, the genetic basis for microbiome assembly in the different ‘spheres of crop plant species is still largely unknown. Therefore, further research should focus on the characterization of microbial communities together with genome-wide association mapping strategies. Genome-wide association studies traditionally rely on plant

populations with hundreds of accessions in order to associate specific regions in the genome with phenotypic traits of interest. Accordingly, studies with a significant number of well characterized RIL populations must be designed and a detailed microbial characterization, taxonomically and functionally, of a specific plant compartment at an appropriate developmental stage must be performed. Ideally, genomic regions associated with the abundance of specific rhizobacterial families or functional microbial genes can be pinpointed in the plant genomes. In this way, microbiome-based breeding programs would include specific genomic regions responsible for the recruitment of specific bacterial groups with beneficial properties, such as diazotrophic bacteria, plant growth promotors or biocontrol agents.

Concluding remarks

The research presented in this thesis indicated that domestication of common bean had a significant effect on the abundance of specific bacterial groups in the rhizosphere and in the spermosphere. These differences may be due to divergent root phenotypic traits and/or different exudation patterns. These genotype-dependent effects were primarily visible in an agricultural soil, where the soil physicochemical conditions and the resident microbial communities were optimal for common bean growth. The higher abundance of Bacteroidetes in the rhizosphere of wild accessions as well as the higher abundance of Actinobacterial families in the rhizosphere and spermosphere of modern accessions were among the specific trends observed in this study. Intriguingly, several families within Bacteroidetes have been also found enriched in and on roots of other wild plant species. What the underlying mechanisms are that drive compositional and functional differences between the microbiomes of wild relatives and modern accessions should be the focus of

future research, including studies addressing the functional impact of these microbiome shifts on plant growth, development and health.

References

Abbo S, Pinhasi van-Oss R, Gopher A, Saranga Y, Ofner I, Peleg Z. 2014. Plant domestication versus crop evolution: a conceptual framework for cereals and grain legumes. *Trends Plant Sci* 19(6):351-60.

Adam E, Bernhart M, Müller H, Winkler J, Berg G. 2018. The *Cucurbita pepo* seed microbiome: genotype-specific composition and implications for breeding. *Plant Soil* 422:35-49.

Aira M, Gómez-Brandón M, Lazcano C, Bååth E, Domínguez J. 2010. Plant genotype strongly modifies the structure and growth of maize rhizosphere microbial communities. *Soil Biol Biochem* 42:2276-2281.

Aguilar OM, Riva O, Peltzer E. 2004. Analysis of *Rhizobium etli* and of its symbiosis with wild *Phaseolus vulgaris* supports coevolution in centres of host diversification. *Proc Natl Acad Sci U S A* 101(37):13548-13553.

Akibode S, Maredia M. 2011. Global and regional trends in production, trade and consumption of food legume crops. (Report submitted to the CGIAR Standing Panel on Impact Assessment). http://www.cgiar.org/our-research/crop-factsheets/beans/

Aleklett K, Hart M. 2013. The root microbiota—a fingerprint in the soil? *Plant Soil* 370:671-686.

Aleklett K, Leff JW, Fierer N, Hart M. 2015. Wild plant species growing closely connected in a subalpine meadow host distinct root-associated bacterial communities. *PeerJ* 3:e804.

Amarger N. 2001. Rhizobia in the field. *Adv Agron* 73:109-168.

Arrieta MC, Stiemsma LT, Amenyogbe N, Brown EM, Finlay B. 2014. The intestinal microbiome in early life: health and disease. *Front Immunol* 5:427.

Badri DV, Chaparro JM, Zhang R, Shen Q, Vivanco JM. 2013. Application of natural blends of phytochemicals derived from the root exudates of Arabidopsis to the soil reveal that phenolic-related compounds predominantly modulate the soil microbiome. *J Biol Chem* 288(7):4502-4512.

Badri DV, Vivanco JM. 2009. Regulation and function of root exudates. *Plant Cell Environ* 32(6):666-81.

Badri DV, Weir TL, van der Lelie D, Vivanco JM. 2009. Rhizosphere chemical dialogues: plant-microbe interactions. *Curr Opin Biotechnol* 20(6):642-50.

Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM. 2006. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu Rev Plant Biol* 57:233-66.

Bakker PAHM, Berendsen RL, Doombos RF, Wintermans PCA, Pieterse CML. 2013. The rhizosphere revisited: root microbiomics. *Front Plant Sci* 4:165.

Bakker MG, Manter DK, Sheflin AM, Weir TL, Vivanco JM. 2012. Harnessing the rhizosphere microbiome through plant breeding and agricultural management. *Plant Soil* 360(1-2):1-13.

Barberán A, Bates ST, Casamayor EO, Fierer N. 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J* 6(2):343-51.

Barret M, Briand M, Bonneau S, Préveaux A, Valière S, Bouchez O, *et al*. 2015. Emergence shapes the structure of the seed microbiota. *Appl Environ Microbiol* 81(4):1257-66.

Bastian M, Heymann S, Jacomy M. 2009. Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media: San Jose, California.

Beckers B, Op De Beeck M, Weyens N, Van Acker R, Van Montagu M, *et al*. 2016. Lignin engineering in field-grown poplar trees affects the endosphere bacterial microbiome. *Proc Natl Acad Sci U S A* 113:2312-7.

Bednarek P, Kwon C, Schulze-Lefert P. 2010. Not a peripheral issue: secretion in plant-microbe interactions. *Curr Opin Plant Biol* 13(4):378-87.

Beebe S, Rengifo J, Gaitan E, Duque MC, Tohme J. 2001. Diversity and origin of Andean landraces of common bean. *Crop Sci* 41(3):854-862.

Berendsen RL, Pieterse CMJ, Bakker PAHM. 2012. The rhizosphere microbiome and plant health. *Trends Plant Sci* 17(8):478-486.

Berendsen RL, Vismans G, Yu K, Song Y, de Jonge R, Burgman WP, *et al*. 2018. Disease-induced assemblage of a plant-beneficial bacterial consortium. *ISME J* 12(6):1496-1507.

Berg G, Grube M, Schloter M, Smalla K. 2014. Unraveling the plant microbiome: looking back and future perspectives. *Front Microbiol* 5:148.

Berg G, Raaijmakers JM. 2018. Saving seed microbiomes. *ISME J* 12(5):1167-1170.

Berg G, Smalla, K. 2009. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiol Ecol* 68: 1-13.

Berlemont R, Martiny AC. 2015. Genomic potential for polysaccharide deconstruction in bacteria. *Appl Environ Microbiol* 81(4):1513-19.

Berry D, Widder S. 2014. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front Microbiol* 5:219.

Bingol K. 2018. Recent advances in targeted and untargeted metabolomics by NMR and MS/NMR methods. *High Throughput* 7(2):E9.

Bitocchi E, Bellucci E, Giardini A, Rau D, Rodriguez M, Biagetti E, *et al*. 2013. Molecular analysis of the parallel domestication of the common bean. *Phaseolus vulgaris*. in Mesoamerica and the Andes. *New Phytol* 197:300–313.

Bitocchi E, Nanni L, Bellucci E, Rossi M, Giardini A, P Spagnoletti Zeuli, *et al*. 2012. Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. *Proc Nat Acad Sci U S A* 109(14): E788-E796.

Blair MW, Soler A, Cortés AJ. 2012. Diversification and population structure in common beans (*Phaseolus vulgaris* L). *PLoS One* 7(11):e49488.

Blaser MJ. 2018. The past and future biology of the human microbiome in an age of extinctions. *Cell* 172:1173-1177.

Boon E, Meehan CJ, Whidden C, Wong DH, Langille MG, Beiko RG. 2013. Interactions in the microbiome: communities of organisms and communities of genes. *FEMS Microbiol Rev* 38(1):90-118.

Bouffaud M-L, Poirier M-A, Mulle D and Moënne-Loccoz, Y. 2014. Root microbiome relates to plant host evolution in maize and other Poaceae. *Environ Microbiol* 16:2804-2814.

Bouwmeester HJ, Roux C, Lopez-Raez JA, Bécard G. 2007. Rhizosphere communication of plants, parasitic plants and AM fungi. *Trends Plant Sci* 12(5):224-30.

Brewer A, Williamson M. 1994. A new relationship for rarefaction. *Biodivers Conserv* 3:373-379.

Broughton WJ, Hernández G, Blair M, Beebe S, Gepts P, Vanderleyden J. 2003. Beans (*Phaseolus* spp.) - model food legumes. *Plant Soil* 252:55-128.

Brown K, DeCoffe D, Molcan E, Gibson DL. 2012. Diet-induced dysbiosis of the intestinal microbiota and the effects on immunity and disease. *Nutrients* 4(8):1095-1119.

Broz AK, Manter DK, Vivanco JM. 2007. Soil fungal abundance and diversity: another victim of the invasive plant *Centaurea maculosa*. *ISME J* 1(8):763-5.

Budak H, Kantar M, Kurtoglu KY. 2013. Drought tolerance in modern and wild wheat. *Sci World J* ID:548246.

Buée M, de Boer W, Martin F, van Overbeek LS, Jurkevitch E. 2009. The rhizosphere zoo: An overview of plant-associated communities of microorganisms, including phages, bacteria, archaea, and fungi, and some of their structuring factors. *Plant Soil* 321:189-212.

Bulgarelli D, Garrido-Oter R, Münch PC, Weiman A, Dröge J, Pan Y *et al*. 2015. Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* 17: 392–403.

Bulgarelli D, Rott M, Schlaeppi K, van Themaat EVL, Ahmadinejad N, Assenza F *et al*. 2012. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* 488:91–95.

Bulgarelli D, Schlaeppi K, Spaepen S, van Themaat EVL, Schulze-Lefert P. 2013. Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol* 64: 807-38.

Burdman S, Volpin H, Kigel J, Kapulnik Y, Okon Y. 1996. Promotion of nod gene inducers and nodulation in common bean. *Phaseolus vulgaris*. roots inoculated with *Azospirillum brasilense* Cd. *Appl Environ Microbiol* 62(8):3030-3.

Busby PE, Soman C, Wagner MR, Friesen ML, Kremer J, Bennett A, *et al*. 2017. Research priorities for harnessing plant microbiomes in sustainable agriculture. *PLoS Biol* 28:15(3):e2001793.

Busse HJ. 2016. Review of the taxonomy of the genus *Arthrobacter*, emendation of the genus *Arthrobacter sensu lato*, proposal to reclassify selected species of the genus *Arthrobacter* in the novel genera *Glutamicibacter* gen. nov., *Paeniglutamicibacter* gen. nov., *Pseudoglutamicibacter* gen. nov., *Paenarthrobacter* gen. nov. and *Pseudarthrobacter* gen. nov., and emended description of *Arthrobacter roseus*. *Int J Syst Evol Microbiol* 66(1):9-37

Buyer JS, Roberts DP, Russek-Cohen E. 1999. Microbial community structure and function in the spermosphere as affected by soil and seed type. *Can J Microbiol* 45(2): 138-144.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, *et al*. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336.

Cardinale M, Grube M, Erlacher A, Quehenberger J, Berg G. 2015. Bacterial networks and co-occurrence relationships in the lettuce root microbiota. *Environ Microbiol* 17:239-52.

Carrión VJ, Cordovez V, Tyc O, Etalo DW, de Bruijn I, de Jager VCL, *et al*. 2018. Involvement of Burkholderiaceae and sulfurous volatiles in disease-suppressive soils. *ISME J* 12(9):2307-2321

Carvalhais LC, Dennis PG, Fedoseyenko D, Hajirezaei MR, Borriss R *et al*. 2011. Root exudation of sugars, amino acids, and organic acids by maize as affected by nitrogen, phosphorus, potassium, and iron deficiency. *J Plant Nutr Soil Sci* 74:3-11.

Carvalhais LC, Dennis PG, Badri DV, Tyson GW, Vivanco JM, Schenk PM. 2013. Activation of the jasmonic acid plant defence pathway alters the composition of rhizosphere bacterial communities. *PLoS One* 8: e56457.

Cesco S, Neumann G, Tomasi N, Pinton R, Weisskopf L. 2010. Release of plant-borne flavonoids into the rhizosphere and their role in plant nutrition. *Plant Soil* 329:1-25.

CGIAR. 2013. Annual progress report 2013: CGIAR Research Program on Grain Legumes, 66 pp.

Chacón MI, Pickersgill SB, Debouck DG. 2005. Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. *Theor Appl Genet* 110:432–444.

Chacón-Fuentes M, Parra L, Lizama M, Seguel I, Urzúa A, Quiroz A. 2017. Plant flavonoid content modified by domestication. *Environ Entomol* 46:1080-1089.

Chagas FO, Pessotti RC, Caraballo-Rodríguez AM, Pupo MT. 2018. Chemical signaling involved in plant-microbe interactions. *Chem Soc Rev* 47(5):1652-1704.

Chaluvadi S, Bennetzen JL. 2018. Species-associated differences in the below-ground microbiomes of wild and domesticated *Setaria*. *Front Plant Sci* 9:1183.

Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4:7.

Chaparro JM, Badri DV, Vivanco JM. 2014. Rhizosphere microbiome assemblage is affected by plant development. *ISME J* 8:790–803.

Chapelle E, Mendes R, Bakker PA, Raaijmakers JM. 2016. Fungal invasion of the rhizosphere microbiome. *ISME J* 10(1):265-8.

Chen WM, Laevens S, Lee TM, Coenye T, De Vos P, Mergeay M, *et al.* 2001. *Ralstonia taiwanensis* sp. nov., isolated from root nodules of *Mimosa* species and sputum of a cystic fibrosis patient. *Int J Syst Evol Microbiol* 51:1729-35.

Chen WM, James EK, Prescott AR, Kierans M, Sprent JI. 2003. Nodulation of *Mimosa* spp. by the beta-proteobacterium *Ralstonia taiwanensis*. *Mol Plant Microbe Interact* 16(12):1051-61.

Chen YH, Gols R, Benrey B. 2015. Crop domestication and its impact on naturally selected trophic interactions. *Annu Rev Entomol* 60:35-58.

Chen YH, Shapiro LR, Benrey B, Cibrián-Jaramillo A. 2017. Back to the origin: in situ studies are needed to understand selection during crop diversification. *Front Ecol Evol* 5:125.

Chen H, Wu H, Yan B, Zhao H, Liu F, Zhang H, *et al*. 2018. Core microbiome of medicinal plant *Salvia miltiorrhiza* seed: A rich reservoir of beneficial microbes for secondary metabolism? *Int J Mol Sci* 19(3):pii: E672

CIAT-PRG. 2018. http://genebank.ciat.cgiar.org

Clayton JB, Vangay P, Huang H, Ward T, Hillmann BM, Al-Ghalith GA, *et al*. 2016. Captivity humanizes the primate microbiome. *Proc Natl Acad Sci U S A*. 113:10376-81.

Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y *et al*. 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 42: D633-D642.

Coleman-Derr D, Desgarennes D, Fonseca-Garcia C, Gross S, Clingenpeel S, Woyke T, *et al*. 2016. Plant compartment and biogeography affect microbiome composition in cultivated and native *Agave* species. *New Phytol* 209(2):798-811.

Comas LH, Becker SR, Cruz VMV, Byrne PF, Dierig DA. 2013. Root traits contributing to plant productivity under drought. *Front Plant Sci* 4:442.

Cook RJ, Thomashow LS, Weller DM, Fujimoto D, Mazzola M, *et al*. 1995. Molecular mechanisms of defense by rhizobacteria against root disease. *Proc Natl Acad Sci U S A* 92(10):4197–4201.

Cortés AJ, Chavarro MC, Madriñán S, This D, Blair MW. 2012. Molecular ecology and selection in the drought-related *Asr* gene polymorphisms in wild and cultivated common bean (*Phaseolus vulgaris* L). *BMC Genet* 13:58.

Coskun D, Britto DT, Shi W, Kronzucker HJ. 2017. How plant root exudates shape the nitrogen cycle. *Trends Plant Sci* 22(8):661-673.

Crittenden AN, Schnorr SL. 2017. Current views on hunter-gatherer nutrition and the evolution of the human diet. *Am J Phys Anthropol* 63:84-109.

Cui YS, Im WT, Yin CR, Lee JS, Lee KC, Lee ST. 2007. *Aeromicrobium panaciterrae* sp. nov., isolated from soil of a ginseng field in South Korea. *Int J Syst Evol Microbiol* 57, 687-91.

Dall'Agnol RF, Plotegher F, Souza RC, Mendes IC, Dos Reis Junior FB, Béna G, *et al*. 2016. *Paraburkholderia nodosa* is the main N2-fixing species trapped by promiscuous common bean (*Phaseolus vulgaris* L.) in the Brazilian 'Cerradão'. *FEMS Microbiol Ecol* 92(8):fiw108.

de Bruijn I, Cheng X, de Jager V, Expósito RG, Watrous J, Patel N, *et al.* 2015. Comparative genomics and metabolic profiling of the genus *Lysobacter*. *BMC Genomics* 16:991.

De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S. *et al*. 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci USA* 107: 14691–96.

de Weert S, Vermeiren H, Mulders IH, Kuiper I, Hendrickx N, Bloemberg GV, *et al*. 2002. Flagella-driven chemotaxis towards exudate components is an important trait for tomato root colonization by *Pseudomonas fluorescens*. *Mol Plant Microbe Interact* 15(11):1173-80.

de Werra P, Huser A, Tabacchi R, Keel C, Maurhofer M. 2011. Plant- and microbe-derived compounds affect the expression of genes encoding antifungal compounds in a pseudomonad with biocontrol activity. *Appl Environ Microbiol* 77(8):2807-12.

Debouck DG, Toro O, Paredes OM, Johnson WC, Gepts P. 1993. Genetic diversity and ecological distribution of *Phaseolus vulgaris* (Fabaceae) in northwestern South America. *Econ Bot* 47:408-423.

Delgado-Baquerizo M, Maestre FT, Reich PB, Trivedi P, Osanai Y, Liu YR, *et al*. 2016. Carbon content and climate variability drive global soil bacterial diversity patterns. *Ecol Monogr* 86(3):373–390.

Desiderio F, Bitocchi E, Bellucci E, Rau D, Rodriguez M, Attene G, *et al.* 2013. Chloroplast microsatellite diversity in *Phaseolus vulgaris*. *Front Plant Sci* 3:1-15.

Dessureault-Rompré J, Nowack B, Schulin R, Luster Jörg. 2006. Modified micro suction cup/rhizobox approach for the in-situ detection of organic acids in rhizosphere soil solution. *Plant Soil* 286:99-107.

Dicke M. 2009. Behavioural and community ecology of plants that cry for help. *Plant Cell Environ* 32(6):654-65.

Dodt M, Roehr JT, Ahmed R, Dieterich C. 2012. FLEXBAR—Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms. *Biology* 1:895–905.

Doebley JF, Gaut BS and Smith BD. 2006. The molecular genetics of crop domestication. *Cell* 127(7):1309-1321.

Dumbrell AJ, Nelson M, Helgason T, Dytham C, Fitter AH. 2010. Relative roles of niche and neutral process in structuring a soil microbial community. *ISME J* 4:337–345.

Earl DA. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4(2):359-361.

Eberl L, Vandamme P. 2016. Members of the genus *Burkholderia*: good and bad guys. *F1000Res* 5(F1000 Faculty Rev):1007

Edgar RC. 2010. Search and clustering hundreds of times faster than BLAST. *Bioinformatics* 26(19):2460-2461.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27(16): 2194-2200.

Edwards J, Johnson C, Santos-Medellín C, Lurie E, Podishetty NK, Bhatnagar S, *et al*. 2015. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci U S A* 112:E911-20.

Eisenhauer N, Lanoue A, Strecker T, Scheu S, Steinauer K, Thakur MP, *et al*. 2017. Root biomass and exudates link plant diversity with soil bacterial and fungal biomass. *Sci Rep* 7:44641.

Elliott GN, Chou JH, Chen WM, Bloemberg GV, Bontemps C, Martínez-Romero E. 2009. *Burkholderia* spp. are the most competitive symbionts of *Mimosa*, particularly under N-limited conditions. *Environ Microbiol* 11(4):762-78.

Ellouze W, Hamel C, Singh AK, Mishra V, DePauw RM, Knox RE. 2018. Abundance of the arbuscular mycorrhizal fungal taxa associated with the roots and rhizosphere soil of different durum wheat cultivars in the Canadian prairies. *Can J Microbiol* 64(8):527-536.

Engelhard M, Hurek T, Reinhold-Hurek B. 2000. Preferential occurrence of diazotrophic endophytes, *Azoarcus* spp., in wild rice species and land races of *Oryza sativa* in comparison with modern races. *Environ Microbiol* 2(2):131-41.

Etienne RS. 2005. A new sampling formula for neutral biodiversity. *Ecol Lett* 8:253-260.

Evanno G, Regnaut S, Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14(8): 2611-2620.

Falagán C, Foesel B, Johnson B. 2017. *Acidicapsa ferrireducens* sp. nov., *Acidicapsa acidiphila* sp. nov., and *Granulicella acidiphila* sp. nov.: novel acidobacteria isolated from metal-rich acidic waters. *Extremophiles.* 21(3):459-469.

Faust K, Raes J. 2012. Microbial interactions: from networks to models. *Nat Rev Microbiol* 10(8):538-50.

Fierer N. 2017. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat Rev Microbiol* 15(10):579-590.

Fierer N, Bradford MA, Jackson RB. 2007. Toward an ecological classification of soil bacteria. *Ecology* 88(6):1354-64.

Fierer N, Jackson RB. 2006. The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A* 103(3):626-31.

Fierer N, Ladau J, Clemente JC, Leff JW, Owens SM, Pollard KS, *et al*. 2013. Reconstructing the microbial diversity and function of pre-agricultural tallgrass prairie soils in the United States. *Science* 342(6158):621-4.

Ferguson, B.J, Lin MH, Gresshoff PM. 2013. Regulation of legume nodulation by acidic growth conditions. *Plant Signal Behav* 8(3):e23426.

Flint-Garcia SA. 2013. Genetics and consequences of crop domestication. *J Agric Food Chem* 61(35):8267-76

Friedman, J. & Alm, E. J. 2012. Inferring correlation networks from genomic survey data. *PLoS Computational Biology.* 8:e1002687.

Garbeva P, van Veen JA, van Elsas JD. 2004. Microbial diversity in soil: Selection of microbial populations by plant and soil type and implications for disease suppressiveness. *Annu Rev Phytopathol* 42:243-70.

García-Palacios P, Milla R, Delgado-Baquerizo M, Martín-Robles N, Alvaro-Sánchez M, Wall DH. 2013. Side-effects of plant domestication: ecosystem impacts of changes in litter quality. *New Phytol* 198(2):504-13.

Gepts P. 2004. Crop domestication as a long-term selection experiment. *Plant Breed Rev* 24(2):1-44.

Gepts P. 1998. Origin and evolution of common bean: past events and recent trends. *HortScience* 33(7):1124-1130.

Gepts P, Bliss FA. 1985. F1 hybrid weakness in the common bean. *J Hered* 76:447-450.

Gepts P, Bliss FA. 1986. Phaseolin variability among wild and cultivated common beans (*Phaseolus vulgaris*) from Colombia. *Econ Bot* 40(4):469-478.

Gepts P, Debouck D. 1991. Origin, domestication and evolution of the common bean (*Phaseolus vulgaris* L.). In: Shoonhoven A van and Voysest O. eds. Common Beans Research for Crop Improvement. CIAT and CAB International, Wallingford, UK. pp. 27-40.

Germida JJ, Siciliano SD. 2001. Taxonomic diversity of bacteria associated with the roots of modern, recent and ancient wheat cultivars. *Biol Fertil Soils* 33: 410–415.

Gomes NCM, Heuer H, Schonfeld J, Costa R, Mendonca-Hagler L, Smalla K. 2001. Bacterial diversity of the rhizosphere of maize (*Zea mays*) grown in tropical soil studied by temperature gradient gel electrophoresis. *Plant Soil* 232:167–80.

Gomez A, Petrzelkova KJ, Burns MB, Yeoman CJ, Amato KR, Vickova K, *et al*. 2016. Gut microbiome of coexisting BaAka Pygmies and Bantu reflects gradients of traditional subsistence patterns. *Cell Rep* 14:2142-53.

Gómez-Expósito R, Postma J, Raaijmakers JM, de Bruijn I. 2015. Diversity and activity of *Lysobacter* species from disease suppressive soils. *Front Microbiol* 6:1243.

González I, Déjean S. 2012. CCA: Canonical correlation analysis. R package version 1.2. http://CRAN.R-project.org/package=CCA.

Gotelli NJ, Colwell RK. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol Lett* 4:379-391.

Gross BL, Olsen KM. 2010. Genetic perspectives on crop domestication. *Trends Plant Sci* 15(9):529-537.

Haichar FZ, Marol C, Berge O, Rangel-Castro JI, Prosser JI, Balesdent J, *et al*. 2008. Plant host habitat and root exudates shape soil bacterial community structure. *ISME J* 2(12):1221-1230.

Haichar FZ, Santaella C, Heulin T and Achouak W. 2014. Root exudates mediated interactions belowground. *Soil Bio and Biochem* 77:69-80.

Hajjar R, Hodgkin T. 2007. The use of wild relatives in crop improvement: A survey of developments over the last 20 years. *Euphytica* 156:1-13.

Hamonts K, Trivedi P, Garg A, Janitz C, Grinyer J, Holford P, *et al*. 2018. Field study reveals core plant microbiota and relative importance of their drivers. *Environ Microbiol* 20(1):124-140.

Hartman K, van der Heijden MG, Roussely-Provent V, Walser JC, Schlaeppi K. 2017. Deciphering composition and function of the root microbiome of a legume plant. *Microbiome* 5:2.

Hassan S, Mathesius U. 2012. The role of flavonoids in root-rhizosphere signalling: opportunities and challenges for improving plant-microbe interactions. *J Exp Bot* 63(9):3429-44.

Hassani MA, Durán P, Hacquard S. 2018. Microbial interactions within the plant holobiont. *Microbiome* 6:58.

Haudry A, Cenci A, Ravel C, Bataillon T, Brunel D, Poncet C, *et al*. 2007. Grinding up wheat: a massive loss of nucleotide diversity since domestication. *Mol Biol Evol* 24(7):1506-1517.

Hayward AC, Fegan N, Fegan M, Stirling GR. 2010. *Stenotrophomonas* and *Lysobacter*: ubiquitous plant-associated gamma-proteobacteria of developing significance in applied microbiology. *J Appl Microbiol* 108(3):756-70.

Henry S, Texier S, Hallet S, Bru D, Dambreville C, Chèneby D, *et al*. 2008. Disentangling the rhizosphere effect on nitrate reducers and denitrifiers: insight into the role of root exudates. *Environ Microbiol* 10(11):3082-92.

Herrera-Flores TS, Acosta-Gallegos JA. 2008. Seed yield of three types of crosses between wild by domesticated genotypes in common bean (*Phaseolus vulgaris* L). *Agricultura Técnica en México* 34(2):167-176.

Hetrick BAD, Wilson GWT, Cox TS. 1992. Mycorrhizal dependence of modern wheat varieties, landraces, and ancestors. *Can J Bot* 70(10):2032-2040.

Hetrick BAD, Wilson GWT, Cox TS. 1993. Mycorrhizal dependence of modern wheat cultivars and ancestors: a synthesis. *Can J Bot* 71(3):512-518.

Hetrick BAD, Wilson GWT, Gill BS, Cox TS. 1995. Chromosome location of mycorrhizal responsive genes in wheat. *Can J Bot* 73:891-897.

Ho A, Di Lonardo DP, Bodelier PLE. 2017. Revisiting life strategy concepts in environmental microbial ecology. *FEMS Microbiol Ecol* 93(3): fix006.

Horton MW, Bodenhausen N, Beilsmith K, Meng D, Muegge BD, Subramanian S, *et al*. 2014. Genome-wide association study of Arabidopsis thaliana leaf microbial community. *Nat Commun* 5:5320.

Iannucci A, Fragasso M, Beleggia R, Nigro F, Papa R. 2017. Evolution of the crop rhizosphere: impact of domestication on root exudates in tetraploid wheat (*Triticum turgidum* L.). *Front Plant Sci* 8:2124.

Illumina. 2013. 16S Metagenomic sequencing library preparation. http://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html

Inceoglu Ö, Abu Al-Soud W, Salles JF, Semenov AV and van Elsas JD. 2011. Comparative analysis of bacterial communities in a potato field as determined by pyrosequencing. *PLoS One* 6(8):e23321.

Inceoglu Ö, Falcão Salles J, van Overbeek L and van Elsas JD. 2010. Effects of plant genotype and growth stage on the betaproteobacterial communities associated with different potato cultivars in two fields. *Appl Environ Microbiol* 76(11):3675-3684.

Jackson, LE. 1995. Root architecture in cultivated and wild lettuce (*Lactuca* spp). *Plant Cell Environ* 18:885–894.

Jabot F, Etienne RF, Chave J. 2008. Reconciling neutral community models and environmental filtering: theory and an empirical test. *Oikos* 117:1308–1320.

Jaccoud D, Peng K, Feinstein D, Kilian A. 2001. Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res* 29(4):e25.

Jeudy C, Adrian M, Baussard C, Bernard C, Bernaud E, Bourion V, *et al*. 2016. RhizoTubes as a new tool for high throughput imaging of plant root development and architecture: test, comparison with pot grown plants and validation. *Plant Methods* 12:31.

Jin J, Wang GH, Liu XB, Liu JD, Chen XL and Herbert SJ. 2009. Temporal and spatial dynamics of bacterial community in the rhizosphere of soybean genotypes grown in a black soil. *Pedosphere* 19(6):808-816.

Johnson WC, Jackson LE, Ochoa O, van Wijk R, Peleman J, St.Clair DA, Michelmore RW. 2000. Lettuce, a shallow-rooted crop, and *Lactuca serriola*, its wild progenitor, differ at QTL determining root architecture and deep soil water exploitation. *Theor Appl Genet* 101:1066–1073.

Jones DL, Hodge A and Kuzyakov Y. 2004. Plant and mycorrhizal regulation of rhizodeposition. *New Phytol* 163:459–480.

Jones DL, Nguyen C, Finlay RD. 2009. Carbon flow in the rhizosphere: carbon trading at the soil–root interface. *Plant Soil* 321:5–33.

Jones RT, Robeson MS, Lauber CL, Hamady M, Knight R, Fierer N. 2009. A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *ISME J* 3(4):442-53.

Joshi NA, Fass JN. 2011. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. Version 1.33. Available at https://github.com/najoshi/sickle

Jousset A, Rochat L, Lanoue A, Bonkowski M, Keel C, Scheu S. 2011. Plants respond to pathogen infection by enhancing the antifungal gene expression of root-associated bacteria. *Mol Plant Microbe Interact* 24(3):352-8.

Kageyama K, Nelson EB. 2003. Differential inactivation of seed exudate stimulation of Pythium ultimum sporangium germination by Enterobacter cloacae influences biological control efficacy on different plant species. *Appl Environ Microbiol* 69(2):1114-20

Kapulnik Y, Kushnir U. 1991. Growth dependency of wild, primitive and modern cultivated wheat lines on vesicular-arbuscular mycorrhiza fungi. *Euphytica* 56:27-36.

Kawasaki A, Donn S, Ryan PR, Mathesius U, Devilla R, Jones A, *et al*. 2016. Microbiome and exudates of the root and rhizosphere of *Brachypodium distachyon*, a model for wheat. *PLoS One* 11(10):e0164533.

Kielak AM, Barreto CC, Kowalchuk GA, van Veen JA, Kuramae EE. 2016. The ecology of acidobacteria: Moving beyond genes and genomes. *Front Microbiol* 7:744.

Kiers ET, Denison RF. 2008. Sanctions, cooperation, and the stability of plant-rhizosphere mutualisms. *Annu Rev Ecol Evol Syst* 39:215–236.

Kiers ET, Hutton MG, Denison RF. 2007. Human selection and the relaxation of legume defences against ineffective rhizobia. *Proc Biol Sci* 274(1629):3119-26.

Kim DH, Kaashyap M, Rathore A, Das RR, Parupalli S, Upadhyaya HD, *et al*. 2014. Phylogenetic diversity of *Mesorhizobium* in chickpea. *J Biosci* 39(3):513-7.

Klaedtke S, Jacques MA, Raggi L, Préveaux A, Bonneau S, Negri V, *et al.* 2016. Terroir is a key driver of seed-associated microbial assemblages. *Environ Microbiol* 18(6):1792-804.

Koide R, Li M, Lewis J, Irby C. 1988. Role of mycorrhizal infection in the growth and reproduction of wild vs. cultivated plants. I. Wild vs. cultivated oats. *Oecologia* 77:537-543.

Köllner TG, Held M, Lenk C, Hiltpold I, Turlings TC, Gershenzon J, Degenhardt J. 2008. A maize (E)-beta-caryophyllene synthase implicated in indirect defense responses against herbivores is not expressed in most American maize varieties. *Plant Cell* 20(2):482-94.

Kolton M, Green SJ, Harel YM, Sela N, Elad Y, Cytryn E. 2012. Draft genome sequence of *Flavobacterium* sp. strain F52, isolated from the rhizosphere of bell pepper (*Capsicum annuum* L. cv. Maccabi). *J Bacteriol* 194:5462-63.

Koomnok C, Teaumroong N, Rerkasem B, Lumyong S. 2007. Diazotroph endophytic bacteria in cultivated and wild rice in Thailand. *ScienceAsia* 33:429-435.

Köster J, Rahmann S. 2012. Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics* 28(19): 2520–2522.

Kraft P, Zeggini E, Ioannidis JPA. 2009. Replication in genome-wide association studies. *Stat Sci* 24: 561–573.

Kramer-Walter KR, Bellingham PJ, Millar TR, Smissen RD, Richardson SJ, Laughlin DC. 2016. Root traits are multidimensional: specific root length is independent from root tissue density and the plant economic spectrum. *J Ecol* 104:1299-1310.

Kroll S, Agler MT, Kemen E. 2017. Genomic dissection of host-microbe and microbe-microbe interactions for advanced plant breeding. *Curr Opin Plant Biol* 36:71-78.

Kuczynski J, Stombaugh J, Walters WA, González A, Caporaso JG, Knight R. 2012. Using QIIME to analyze 16S rRNA gene sequences from microbial communities. *Curr Protoc Chem Biol* 1E:5.1–20.

Kulichevskaya IS, Kostina LA, Valásková V, Rijpstra WI, Damsté JS, de Boer W, *et al*. 2012. *Acidicapsa borealis* gen. nov., sp. nov. and *Acidicapsa ligni* sp. nov., subdivision 1 *Acidobacteria* from *Sphagnum* peat and decaying wood. *Int J Syst Evol Microbiol* 62:1512-20.

Kuramae EE, Yergeau E, Wong LC, Pijl AS, van Veen JA, Kowalchuk GA. 2012. Soil characteristics more strongly influence soil bacterial communities than land-use type. *FEMS Microbiol Ecol* 79(1):12-24.

Kwak YS, Bonsall RF, Okubara PA, Paulitz TC, Thomashow LS, *et al*. 2012. Factors impacting the activity of 2, 4-diacetylphloroglucinol-producing *Pseudomonas fluorescens* against take-all of wheat. *Soil Biol Biochem* 54:48–56.

Lakshmanan V, Kitto SL, Caplan JL, Hsueh YH, Kearns DB, Wu YS *et al*. 2012. Microbe-associated molecular patterns-triggered root responses mediate beneficial rhizobacterial recruitment in *Arabidopsis*. *Plant Physiol* 160:1642-1661.

Lakshmanan V, Selvaraj G, Bais HP. 2014. Functional soil microbiome: belowground solutions to an aboveground problem. *Plant Physiol* 166(2):689-700.

Lanfranco L, Fiorilli V, Venice F, Bonfante P. 2018. Strigolactones cross the kingdoms: plants, fungi, and bacteria in the arbuscular mycorrhizal symbiosis. *J Exp Bot* 69(9):2175-2188.

Lareen A, Burton F, Schäfer P. 2016. Plant root-microbe communication in shaping root microbiomes. *Plant Mol Bio* 90(6):575-87.

Larsson J. 2017. eulerr: Area-Proportional Euler and Venn Diagrams with Circles or Ellipses R package version 3.0.1. https://cran.r-project.org/package=eulerr.

Lauber CL, Hamady M, Knight R, Fierer N. 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl Environ Microbiol* 75(15):5111-20.

Lebeis SL, Paredes SH, Lundberg DS, Breakfield N, Gehring J, McDonald M, *et al*. 2015. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* 349:860-864.

Leff JW, Lynch RC, Kane NC, Fierer N. 2017. Plant domestication and the assembly of bacterial and fungal communities associated with strains of the common sunflower, *Helianthus annuus*. *New Phytol* 214(1):412-423.

Lehmann A, Barto EK, Powell JR, Rillig MC. 2012. Mycorrhizal responsiveness trends in annual crop plants and their wild relatives-a meta-analysis on studies from 1981 to 2010. *Plant Soil* 355:231–250.

Lemaire B, Chimphango SBM, Stirton C, Rafudeen S, Honnay O, Smets E. *et al.* 2016. Biogeographical patterns of legume-nodulating *Burkholderia* spp.: from African fynbos to continental scales. *Appl Environ Microbiol* 82(17), 5099-115.

Lemanceau P, Barret M, Mazurier S, Mondy S, Pivato B, Fort T, *et al*. 2017. Plant communication with associated microbiota in the spermosphere, rhizosphere and phyllosphere. *Adv Bot Res* 82:101-133.

Lemanceau, P., Blouin, M., Muller, D. & Moënne-Loccoz, Y. 2017. Let the core microbiota be functional. *Trends in Plant Science.* 22(7), 583-595.

Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 44(Web Server issue):W242–W245.

Levy A, Salas-Gonzalez I, Mittelviefhaus M, Clingenpeel S, Herrera Paredes S, Miao J, *et al*. Genomic features of bacterial adaptation to plants. *Nat Genet* 50(1):138-150.

Ley RE, Peterson DA, Gordon JI. 2006. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124(4):837-48.

Ley RE, Turnbaugh PJ, Klein S, Gordon JI. 2006. Microbial ecology: human gut microbes associated with obesity. *Nature* 444:1022-23.

Li M, Li D, Tang Y, Wu F, Wang J. 2017. CytoCluster: A Cytoscape Plugin for Cluster Analysis and Visualization of Biological Networks. *Int J Mol Sci* 18(9):1880.

Ling N, Raza W, Ma J, Huang Q, Shen Q. 2011. Identification and role of organic acids in watermelon root exudates for recruiting *Paenibacillus polymyxa* SQR-21 in the rhizosphere. *Eur J Soil Biol* 47:374–379.

Liu R, Hong J, Xu X, Feng Q, Zhang D, Gu YR, *et al*. 2017. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention *Nat Med* 23:859-68.

Liu CW, Murray JD. 2016. The role of flavonoids in nodulation host-range specificity: An update. *Plants* 5(3):33.

López-López A, Rogel MA, Ormeño-Orrillo E, Martínez-Romero J, Martínez-Romero E. 2010. *Phaseolus vulgaris* seed-borne endophytic community with novel bacterial species such as *Rhizobium endophyticum* sp. nov. *Syst Appl Microbiol* 33(6):322-7.

López-Velasco G, Carder PA, Welbaum GE, Ponder MA. 2013. Diversity of the spinach (*Spinacia oleracea*) spermosphere and phyllosphere bacterial communities. *FEMS Microbiol Lett* 346(2):146-54

Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S *et al*. 2012. Defining the core *Arabidopsis thaliana* root microbiome. *Nature* 488: 86-90.

Manteca A, Sanchez J. 2009. *Streptomyces* development in colonies and soils. *Appl Environ Microbiol* 75(9):2920-4.

Marques JM, da Silva TF, Vollu RE, Blank AF, Ding GC, Seldin L, *et al*. 2014. Plant age and genotype affect the bacterial community composition in the tuber rhizosphere of field-grown sweet potato plants. *FEMS Microbiol Ecol* 88(2):424-35.

Marschner P, Neumann G, Kania A, Weiskopf L, Lieberei R. 2002. Spatial and temporal dynamics of the microbial community structure in the rhizosphere of cluster roots of white lupin (*Lupinus albus* L.). *Plant Soil* 246: 167-164.

Martin FM, Uroz S, Barker DG. 2017. Ancestral alliances: Plant mutualistic symbioses with fungi and bacteria. *Science* 356(6340):eaad4501.

Martín-Robles N, Morente-López J, Milla R. 2015. Digging up the roots of crop evolution. Libro de resúmenes de comunicaciones del 4º Congreso Ibérico de Ecología, Coimbra, Portugal. doi: 10.7818/4IberianEcologicalCongress.

Martínez-Romero E, Segovia L, Mercante FM, Franco AA, Graham P, Pardo MA. 1991. *Rhizobium tropici*, a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena* sp. trees. *Int J Syst Bacteriol* 41(3):417-26.

Matson PA, Parton WJ, Power AG, Swift MJ. 1997. Agricultural intensification and ecosystem properties. *Science* 277(5325):504-509.

Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD. 2012. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics* 13: 31.

Mazzola M, Funnell DL, Raaijmakers JM. 2004. Wheat cultivar-specific selection of 2,4-diacetylphloroglucinol-producing fluorescent pseudomonas species from resident soil populations. *Microbiol Ecol* 48: 338–348.

McDonald D, Clemente JC, Kuczynski J, Rideout JR, Stombaugh J, Wendel D *et al*. 2012. The Biological Observation Matrix. BIOM. format or: how I learned to stop worrying and love the ome-ome. *Gigascience* 1: 7.

McKenzie VJ, Song SJ, Delsuc F, Prest TL, Oliverio AM, Korpita TM, *et al*. 2017. The effects of captivity on the mammalian gut microbiome. *Integr Comp Biol*. 57:690-704.

McMurdie PJ, Holmes S. 2013. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8(4):e61217.

McNear Jr DH. 2013. The Rhizosphere - Roots, soil and everything in between. *Nature Education Knowledge* 4(3):1.

Mendes LW, Kuramae EE, Navarrete AA, van Veen JA, Tsai SM. 2014. Taxonomical and functional microbial community selection in soybean rhizosphere. *ISME J* 8:1577–1587.

Mendes LW, Raaijmakers JM, de Hollander M, Mendes R, Tsai SM. 2018. Influence of resistance breeding in common bean on rhizosphere microbiome composition and function. *ISME J* 12(1):212-224.

Mendes R, Garbeva P, Raaijmakers JM.. 2013. The rhizosphere microbiome: significance of plant beneficial, plant pathogenic and human pathogenic microorganisms. *FEMS Microbiol Rev* 37:634–663.

Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JHM *et al*. 2011. Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* 332 (6033):1097-1100.

Meyer RS, DuVal AE, Jensen HR. 2012. Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. *New Phytol* 196:29-48.

Micallef SA, Channer S, Shiaris MP, Colón-Carmona A. 2009a. Plant age and genotype impact the progression of bacterial community succession in the Arabidopsis rhizosphere. *Plant Signal Behav* 4(8):777-80.

Micallef SA, Shiaris MP and Colón-Carmona A. 2009b. Influence of *Arabidopsis thaliana* accessions on rhizobacterial communities and natural variation in root exudates. *J Exp Bot* 60(6):1729-1742.

Miller ES, Woese CR, Brenner S. 1991. Description of the erythromycin-producing bacterium *Arthrobacter* sp. strain NRRL B-3381 as *Aeromicrobium erythreum* gen. nov., sp. nov. *Int J Syst Bacteriol* 41(3):363-8.

Moe LA. 2013. Amino acids in the rhizosphere: From plants to microbes. *Am J Bot* 100(9):1692–1705.

Mönchgesang S, Strehmel N, Schmidt S, Westphal L, Taruttis F, Müller E, *et al*. 2016. Natural variation of root exudates in *Arabidopsis thaliana*-linking metabolomic and genomic data. *Sci Rep* 6:29033.

Moulin L, Munive A, Dreyfus B, Boivin-Masson C. 2001. Nodulation of legumes by members of the beta-subclass of Proteobacteria. *Nature* 411(6840):948-50.

Mounier E, Hallet S, Chèneby D, Benizri E, Gruet Y, Nguyen C, *et al*. 2004. Influence of maize mucilage on the diversity and activity of the denitrifying community. *Environ Microbiol* 6(3):301-12.

Mueller UG, Sachs JL. 2015. Engineering microbiomes to improve plant and animal health. *Trends Microbiol* 23:606-17.

Mummey DL, Rillig MC. 2006. The invasive plant species *Centaurea maculosa* alters arbuscular mycorrhizal fungal communities in the field. *Plant Soil* 288:81–90.

Mutch LA, Young JP. 2004. Diversity and specificity of *Rhizobium leguminosarum* biovar *viciae* on wild and cultivated legumes. *Mol Ecol* 13(8):2435-44.

Naz AA, Arifuzzaman M, Muzammil S, Pillen K, Léon J. 2014. Wild barley introgression lines revealed novel QTL alleles for root and related shoot traits in the cultivated barley (*Hordeum vulgare* L.) *BMC Genetics* 15:107.

Naz AA, Ehl A, Pillen K, Léon J. 2012. Validation for root-related quantitative trait locus effects of wild origin in the cultivated background of barley (*Hordeum vulgare* L.). *Plant Breed* 3:392-398.

Neal AL, Ahmad S, Gordon-Weeks R, Ton J. 2012. Benzoxazinoids in root exudates of maize attract *Pseudomonas putida* to the rhizosphere. *PLoS One* 7(4):e35498.

Nelson EB. 1990. Exudate molecules initiating fungal responses to seeds and roots. *Plant Soil* 129:61-73.

Nelson EB. 2004. Microbial dynamics and interactions in the spermosphere. *Annu Rev Phytopathol* 42:271-309.

Nelson EB. 2018. The seed microbiome: Origins, interactions, and impacts. *Plant Soil* 422:7-24.

Nevo E, Chen G. 2010. Drought and salt tolerances in wild relatives for wheat and barley improvement. *Plant Cell Environ* 33:670–685.

Ofek M, Hadar Y, Minz D. 2012. Ecology of root colonizing *Massilia* (Oxalobacteraceae). *PLoS One* 7(7):e40117

Ofek M, Voronov-Goldman M, Hadar Y, Minz D.. 2014. Host signature effect on plant root-associated microbiomes revealed through analyses of resident vs. active communities. *Environ Microbiol* 16:2157-2167.

Ofek-Lalzar M, Sela N, Goldman-Voronov M, Green SJ, Hadar Y, *et al*. 2014. Niche and host-associated functional signatures of the root surface microbiome. *Nat commun* 5:4950.

Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson GL, Solymos P *et al*. 2016, 2017. vegan: Community ecology package. R package versions 2.4-0 and 2.4-4. http://CRAN.R-project.org/package=vegan

Olsen KM, Caicedo AL, Polato N, McClung A, McCouch S, Purugganan MD. 2006. Selection under domestication: evidence for a sweep in the rice waxy genomic region. *Genetics* 173(2):975-83.

Palaisa K, Morgante M, Tingey S, Rafalski A. 2004. Long-range patterns of diversity and linkage disequilibrium surrounding the maize Y1 gene are indicative of an asymmetric selective sweep. *Proc Natl Acad Sci USA* 101(26):9885-90.

Parks DH, Tyson GW, Hugenholtz P, Beiko RG. 2014. STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 30(21):3123-4.

Paulson JN, Stine OC, Bravo HC, Pop M. 2013. Differential abundance analysis for microbial marker-gene surveys. *Nat Methods* 10:1200–1202.

Paulson JN, Talukder H, Pop M, Bravo HC. 2016. metagenomeSeq: Statistical analysis for sparse high-throughput sequencing. Bioconductor package: 1.12.0. http://cbcb.umd.edu/software/metagenomeSeq

Peiffer JA, Spor A, Koren O, Jin Z, Tringe SG, Dangl JL *et al*. 2013. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc Natl Acad Sci U S A* 110(16): 6548–6553.

Pérez-Jaramillo JE, Carrión VJ, Bosse M, Ferrão LFV, de Hollander M, Garcia AAF, *et al.* 2017. Linking rhizosphere microbiome composition of wild and domesticated *Phaseolus vulgaris* to genotypic and root phenotypic traits. *ISME J* 11:2244–57.

Pérez-Jaramillo JE, Carrión VJ, de Hollander M, Raaijmakers JM. 2018. The wild side of plant microbiomes. *Microbiome* 6:143.

Pérez-Jaramillo JE, Mendes R, Raaijmakers JM. 2016. Impact of plant domestication on rhizosphere microbiome assembly and functions. *Plant Mol Biol* 90: 635–644.

Petersen DJ, Srinivasan M, Chanway CP. 1996. *Bacillus polymyxa* stimulates increased *Rhizobium etli* populations and nodulation when co-resident in the rhizosphere of *Phaseolus vulgaris. FEMS Microbiol Lett* 142(2-3):271-6.

Pfeiffer S, Mitter B, Oswald A, Schloter-Hai B, Schloter M, Declerck S. 2017.Rhizosphere microbiomes of potato cultivated in the High Andes show stable and dynamic core microbiomes with different responses to plant development. *FEMS Microbiol Ecol* 93(2): fiw242.

Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH. 2013. Going back to the roots: the microbial ecology of the rhizosphere. *Nature Rev Microbiol* 11: 789-799.

Placido DF, Campbell MT, Folsom JJ, Cui X, Kruger GR, Baenziger PS, Harkamal WH. 2013. Introgression of novel traits from a wild wheat relative improves drought adaptation in wheat. *Plant Physiol* 161(4):1806–1819.

Pongsilp N, Nimnoi P and Lumyong S. 2012. Genotypic diversity among rhizospheric bacteria of three legumes assessed by cultivation-dependent and cultivation-independent techniques. *World J Microbiol Biotechnol* 28(2):615-26.

Porch TG, Beaver JS, Debouck DG, Jackson SA, Kelly JD. 2013. Use of wild relatives and closely related species to adapt common bean to climate change. *Agronomy* 3:433-461.

Postma J, Schilder MT, van Hoof RA. 2011. Indigenous populations of three closely related *Lysobacter* spp. in agricultural soils using real-time PCR. *Microb Ecol* 62(4): 948-58.

Price MN, Dehal PS, Arkin AP. 2009. FastTree: Computing large minimum-evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 26:1641-1650.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155(2):945-959.

Puopolo G, Tomada S, Pertot I. 2018. The impact of the omics era on the knowledge and use of *Lysobacter* species to control phytopathogenic micro-organisms. *J Appl Microbiol* 124(1):15-27.

Purugganan MD, Fuller DQ. 2009. The nature of selection during plant domestication. *Nature* 457:843-848.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, *et al*. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Raaijmakers JM, Paulitz TC, Steinberg C, Alabouvette C, Moënne-Loccoz Y. 2009. The rhizosphere: a playground and battlefield for soilborne pathogens and beneficial microorganisms. *Plant Soil* 321:341-361.

Raboin, L-M, Razafimahafaly AHD, Rabenjarisoa MB, Rabary B, Deuserre J, Becquer T. 2016. Improving the fertility of tropical acid soils: liming versus biochar application? A long term comparison in the highlands of Madagascar. *Field Crops Res* 199:99–108.

Ram SG, Thiruvengadam V, Vinod KK. 2007. Genetic diversity among cultivars, landraces and wild relatives of rice as revealed by microsatellite markers. *J Appl Genet* 48(4):337-45.

Ramirez KS, Craine JM, Fierer N. 2012. Consistent effects of nitrogen amendments on soil microbial communities and processes across biomes. *Glob Chang Biol* 18:1918-1927.

Ramirez KS, Knight CG, de Hollander M, Brearley FQ, Constantinides B, *et al*. 2018. Detecting macroecological patterns in bacterial communities across independent studies of global soils. *Nat Microbiol* 3(2):189-196.

Rasmann S, Köllner TG, Degenhardt J, Hiltpold I, Toepfer S, Kuhlmann U, *et al*. 2005. Recruitment of entomopathogenic nematodes by insect-damaged maize roots. *Nature* 434(7034):732-737.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W *et al*. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43(7):e47.

Rodrigues JL, Pellizari VH, Mueller R, Baek K, Jesus Eda C, Paula FS, *et al*. 2013. Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. *Proc Natl Acad Sci U S A* 110(3):988-993.

Rodrigues JL, Pellizari VH, Mueller R, Baek K, Jesus Eda C, Paula FS, *et al.* 2013. Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. *Proc Natl Acad Sci U S A* 110(3):988-993.

Rognes T, Mahé F, Flouris T, Quince C, Nichols B. 2015. VSEARCH version 1.0.16. https://github.com/torognes/vsearch

Rohrbacher F, St-Arnaud M. 2016. Root exudation: The ecological driver of hydrocarbon rhizoremediation. *Agronomy* 6(1):19.

Rudrappa T, Czymmek KJ, Paré PW, Bais HP. 2008. Root-secreted malic acid recruits beneficial soil bacteria. *Plant Physiol* 148(3):1547-56.

Ruiz-Pérez CA, Restrepo S, Zambrano MM. 2016. Microbial and functional diversity within the phyllosphere of Espeletia species in an Andean high-mountain ecosystem. *Appl Environ Microbiol* 82:1807–1817.

Rybakova D, Mancinelli R, Wikström M, Birch-Jensen AS, Postma J, Ehlers RU, *et al.* 2017. The structure of the *Brassica napus* seed microbiome is cultivar-dependent and affects the interactions of symbionts and pathogens. *Microbiome* 5(1):104.

Saleem M, Law AD, Sahib MR, Pervaiz ZH, Zhang Q. 2018. Impact of root system architecture on rhizosphere and root microbiome. *Rhizosphere* 6:47-51.

Sánchez PA, Logan JT. 1992. Myths and science about the chemistry and fertility of soils in the tropics. In: Lal R & Sanchez PA. Eds. Myths and Science of Soils in the Tropics, SSSA Special Publication 29. ASA and SSSA, Madison, WI, USA.

Sangabriel-Conde W, Maldonado-Mendoza IE, Mancera-López ME, Cordero-Ramírez JD, Trejo-Aguilar D, Negrete-Yankelevich S. 2015. Glomeromycota associated with Mexican native maize landraces in Los Tuxtlas, Mexico. *Appl Soil Ecol* 87:63–71.

Sangabriel-Conde W, Negrete-Yankelevich S, Maldonado-Mendoza IE, Trejo-Aguilar D. 2014. Native maize landraces from Los Tuxtlas, Mexico show varying mycorrhizal dependency for P uptake. *Biol Fertil Soils* 50:405–414.

Sasse J, Martinoia E, Northen T. 2018. Feed your friends: Do plant exudates shape the root microbiome? *Trends Plant Sci* 23:25-41.

Schiltz S, Gaillard I, Pawlicki-Jullian N, Thiombiano B, Mesnard F, Gontier E. 2015. A review: what is the spermosphere and how can it be studied? *J Appl Microbiol* 119(6):1467-81.

Schlaeppi K, Dombrowski N, Oter RG, van Themaat EVL, Schulze-Lefert P. 2014. Quantitative divergence of the bacterial root microbiota in *Arabidopsis thaliana* relatives. *Proc Natl Acad Sci U S A* 111:585–592.

Schlub RL, Schmitthenner AF. 1978. Effects of soybean seed coat cracks on seed exudation and seedling quality in soil infested with *Pythium ultimum*. *Phytopathology* 68:1186-1191.

Schmalenbach I, Körber N, Pillen K. 2008. Selecting a set of wild barley introgression lines and verification of QTL effects for resistance to powdery mildew and leaf rust. *Theor Appl Genet* 117(7):1093–1106.

Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, *et al*. 2014. Gut microbiome of the Hadza hunter-gatherers. *Nature Commun* 5:3654.

Schnorr SL, Sankaranarayanan K, Lewis CM Jr, Warinner C. 2016. Insights into human evolution from ancient and contemporary microbiome studies. *Curr Opin Genet Dev* 41:14-26.

Schultze M, Kondorosi A. 1998. Regulation of symbiotic root nodule development. *Annu Rev Genet* 32:33-57.

Schulz-Bohm K, Gerards S, Hundscheid M, Melenhorst J, de Boer W, Garbeva P. 2018. Calling from distance: attraction of soil bacteria by plant root volatiles. *ISME J* 12(5):1252-1262.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, *et al*. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* 13(11):2498-504.

Shaposhnikov AI, Morgounov AI, Akin B, Makarova NM, Belimov AA, Tikhonovich IA. Comparative characteristics of root systems and root exudation of synthetic, landrace and modern wheat varieties. *Agricultural Biology*. 2016;51:68-78.

Shenton M, Iwamoto C, Kurata N, Ikeo K. 2016. Effect of wild and cultivated rice genotypes on rhizosphere bacterial community composition. *Rice* 9(1):42.

Shi S, Nuccio E, Herman DJ, Rijkers R, Estera K, Li J, *et al*. 2015. Successional trajectories of rhizosphere bacterial communities over consecutive seasons. *mBio* 6:e00746.

Shi S, Richardson AE, O'Callaghan M, DeAngelis KM, Jones EE, Stewart A, *et al*. 2011. Effects of selected root exudate components on soil bacterial communities. *FEMS Microbiol Ecol* 77(3):600-10.

Short GE, Lacy ML. 1976. Carbohydrate exudation from pea seeds: Effect of cultivar, seed age, seed color, and temperature. *Phytopathology* 66:182-187.

Simko I, Atallah AJ, Ochoa OE, Antonise R, Galeano CH, Truco MJ, Michelmore MW. 2013. Identification of QTLs conferring resistance to downy mildew in legacy cultivars of lettuce. *Sci* Rep 3:2875.

Simon HM, Smith KP, Dodsworth JA, Guenthner B, Handelsman J, Goodman RM. 2001. Influence of tomato genotype on growth of inoculated and indigenous bacteria in the spermosphere. *Appl Environ Microbiol* 67(2):514-20.

Singh SP. 2001. Broadening the genetic base of common bean cultivars: a review. *Crop Sci* 41:1659–1675.

Singh P, Mehrotra RS. 1980. The influence of cultivar and temperature on carbohydrate and amino acid exudation from gram seeds and on pre-emergence damping-off by *Rhizoctonia bataticola*. *Plant Soil* 55:261-268.

Smalla K, Wieland G, Buchner A, Zock A, Parzy J, Kaiser S, Roskot N, Heuer H and Berg G. 2001. Bulk and rhizosphere soil bacterial communities studied by denaturing gradient gel electrophoresis: plant-dependent enrichment and seasonal shifts revealed. *Appl Environ Microbiol* 67(10):4742-4751.

Smith KP, Handelsman J, Goodman RM. 1997. Modeling dose-response relationships in biological control: Partitioning host responses to the pathogen and biocontrol agent. *Phytopathology* 87(7):720-729.

Smith KP, Goodman RM. 1999. Host variation for interactions with beneficial plant-associated microbes. *Annu Rev Phytopathol* 37:473-491.

Smýkal P, Nelson MN, Berger JD, von Wettberg EJ. 2018. The impact of genetic changes during crop domestication. *Agronomy* 8:119.

Sonnante G, Stockton T, Nodari RO, Becerra-Velásquez VL, Gepts P. 1994. Evolution of genetic diversity during the domestication of common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet* 89:629-635.

Stopnisek N, Bodenhausen N, Frey B, Fierer N, Eberl L, Weisskopf L. 2014. Genus-wide acid tolerance accounts for the biogeographical distribution of soil *Burkholderia* populations. *Environ Microbiol* 16(6):1503-12.

Stringlis IA, Yu K, Feussner K, de Jonge R, Van Bentum S, Van Verk MC, *et al*. 2018. MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. *Proc Natl Acad Sci U S A* 115(22):E5213-E5222.

Sugiyama A, Ueda Y, Takase H and Yazaki K. 2014a. Pyrosequencing assessment of rhizosphere fungal communities from a soybean field. *Can J Microbiol* 60(10):687-690.

Sugiyama A, Ueda Y, Zushi T, Takase H and Yazaki K. 2014b. Changes in the bacterial community of soybean rhizospheres during growth in the field. *PLoS One* 9(6): e100709.

Szoboszlay M, Lambers J, Chappell J, Kupper JV, Moe LA, McNear Jr DH. 2015. Comparison of root system architecture and rhizosphere microbial communities of Balsas teosinte and domesticated corn cultivars. *Soil Bio and Biochem* 80:34-44.

Talbi C, Delgado MJ, Girard L, Ramírez-Trujillo A, Caballero-Mellado J, Bedmar EJ. 2010. *Burkholderia phymatum* strains capable of nodulating *Phaseolus vulgaris* are present in Moroccan soils. *Appl Environ Microbiol* 76(13):4587-91.

Tardy V, Chabbi A, Charrier X, de Berranger C, Reignier T, Dequiedt S, *et al.* 2015. Land use history shifts in situ fungal and bacterial successions following wheat straw input into the soil. *PLoS One* 10(6):e0130672.

Thomas F, Hehemann JH, Rebuffet E, Czjzek M, Michel G. 2011. Environmental and gut bacteroidetes: The food connection. *Front Microbiol* 2:93.

Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, *et al*. 2017. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 551(7681):457-463.

Tian L, Zhou X, Ma L, Xu S, Nasir F, Tian C. 2017. Root-associated bacterial diversities of *Oryza rufipogon* and *Oryza sativa* and their influencing environmental factors. *Arch Microbiol* 199(4):563-571.

Toju H, Peay KG, Yamamichi M, Narisawa K, Hiruma K, Naito K, *et al*. 2018. Core microbiomes for sustainable agroecosystems. *Nat Plants* 4(5):247-257.

Toro O, Thome J, Debouck DG. 1990. Wild bean (*Phaseolus vulgaris* L.): Description and distribution. Centro Internacional de Agricultura Tropical (CIAT); International Board for Plant Genetic Resources (IBPGR), Cali, Colombia. 106 pp. ISBN: 958-9183-22-0.

Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444(7122):1027-1031.

Turner TR, James EK, Poole PS. 2013a. The plant microbiome. *Genome Biol* 14:209.

Turner TR, Ramakrishnan K, Walshaw J, Heavens D, Alston M, Swarbreck D, *et al*. 2013b. Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *ISME J* 7(12):2248-2258.

Uksa M, Schloter M, Endesfelder D, Kublik S, Engel M, Kautz T, *et al*. 2015. Prokaryotes in subsoil-evidence for a strong spatial separation of different phyla by analysing co-occurrence networks. *Front Microbiol* 6:1269.

Vandenkoornhuyse P, Quaiser A, Duhamel M, Le Van A, Dufresne A. 2015. The importance of the microbiome of the plant holobiont. *New Phytol* 206(4):1196-206.

Viaene T, Langendries S, Beirinckx S, Maes M, Goormachtig S. 2016. Streptomyces as a plant's best friend? *FEMS Microbiol Ecol* 92(8):pii: fiw119

Wallenstein MD. 2017. Managing and manipulating the rhizosphere microbiome for plant health: A systems approach. *Rhizosphere* 3:230-232.

Walters WA, Jin Z, Youngblut N, Wallace JG, Sutter J, Zhang W, *et al*. 2018. Large-scale replicated field study of maize rhizosphere identifies heritable microbes. *Proc Natl Acad Sci U S A* 115(28):7368-7373.

Wang E, Schornack S, Marsh JF, Gobbato E, Schwessinger B, Eastmond P, *et al*. 2012. A common signaling process that promotes mycorrhizal and oomycete colonization of plants. *Curr Biol* 22(23):2242-6.

Wang L, Cao Y, Wang ET, Qiao YJ, Jiao S, Liu ZS, *et al*. 2016. Biodiversity and biogeography of rhizobia associated with common bean (*Phaseolus vulgaris* L.) in Shaanxi Province. *Syst Appl Microbiol* 39(3):211-219.

Wang ZH, Fang H, Chen M. 2017. Effects of root exudates of woody species on the soil anti-erodibility in the rhizosphere in a karst region, China. *PeerJ* 5:e3029.

Ward NL, Challacombe JF, Janssen PH, Henrissat B, Coutinho PM, Wu M *et al.* 2009. Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. *Appl Environ Microbiol* 75(7):2046-56.

Waters MT, Gutjahr C, Bennett T, Nelson DC. 2017. Strigolactone signaling and evolution. *Annu Rev Plant Biol* 68:291-322.

Wickham H. 2009. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Weinert N, Piceno Y, Ding GC, Meincke R, Heuer H, Berg G, *et al*. 2011. PhyloChip hybridization uncovered an enormous bacterial diversity in the rhizosphere of different potato cultivars: many common and few cultivar-dependent taxa. *FEMS Microbiol Ecol* 75(3):497-506.

Weese DJ, Heath KD, Dentinger BT, Lau JA. 2015. Long-term nitrogen addition causes the evolution of less-cooperative mutualists. *Evolution* 69(3):631-42.

Wemheuer, F, Kaiser K, Karlovsky P, Daniel R, Vidal S, Wemheuer B. 2017. Bacterial endophyte communities of three agricultural important grass species differ in their response towards management regimes. *Sci Rep* 7:40914.

Weston LA, Mathesius U. 2013. Flavonoids: their structure, biosynthesis and role in the rhizosphere, including allelopathy. *J Chem Ecol* 39(2):283-97.

Wickham, H. 2009, ggplot2: elegant graphics for data analysis. New York: Springer-Verlag. 260 pp.

Wieland G, Neumann R, Backhaus H. 2001. Variation of microbial communities in soil, rhizosphere, and rhizoplane in response to crop species, soil type, and crop development. *Appl Environ Microbiol* 67(12):5849-54.

Wilhelm RC, Cardenas E, Leung H, Szeitz A, Jensen LD, Mohn WW. 2017. Long-term enrichment of stress-tolerant cellulolytic soil populations following timber harvesting evidenced by multi-omic stable isotope probing. *Front Microbiol* 8:537.

Wintermans PC, Bakker PA, Pieterse CM. 2016. Natural genetic variation in Arabidopsis for responsiveness to plant growth-promoting rhizobacteria. *Plant Mol Biol* 90(6):623-34.

Wissuwa M, Mazzola M, Picard C. 2009. Novel approaches in plant breeding for rhizosphere related traits. *Plant Soil* 321:409–430.

Wolińska A, Kuźniar A, Zielenkiewicz U, Izak D, Szafranek-Nakonieczna A, Banach A, *et al.* 2017. Bacteroidetes as a sensitive biological indicator of agricultural soil usage revealed by culture-independent approach. *Appl Soil Ecol* 119:128–137.

Xing X, Koch AM, Jones AM, Ragone D, Murch S, Hart MM. 2012. Mutualism breakdown in breadfruit domestication. *Proc Biol Sci* 22:279(1731):1122-30.

Xu Y, Wang G, Jin J, Liu J, Zhang Q, Liu Xiaobing. 2009. Bacterial communities in soybean rhizosphere in response to soil type, soybean genotype, and their growth stage. *Soil Biol Biochem* 41:919–925.

Yeoh YK, Dennis PG, Paungfoo-Lonhienne C, Weber L, Brackin R, Ragan MA, *et al.* 2017. Evolutionary conservation of a core root microbiome across plant phyla along a tropical soil chronosequence. *Nat Commun* 8(1):215.

Yeoh YK, Paungfoo-Lonhienne C, Dennis PG, Robinson N, Ragan MA, Schmidt S, *et al.* 2016. The core root microbiome of sugarcanes cultivated under varying nitrogen fertilizer application. *Environ Microbiol* 18(5):1338-51.

Yin C, Hulbert SH, Schroeder KL, Mavrodi O, Mavrodi D, Dhingra A, *et al*. 2013. Role of bacterial communities in the natural suppression of *Rhizoctonia solani* bare patch disease of wheat (*Triticum aestivum* L.). *Appl Environ Microbiol* 79:7428-7438.

Yoon JH, Lee CH, Oh TK. 2005. *Aeromicrobium alkaliterrae* sp. nov., isolated from an alkaline soil, and emended description of the genus *Aeromicrobium*. *Int J Syst Evol Microbiol* 55:2171-5.

Zachow C, Müller H, Tilcher R, Berg G. 2014. Differences between the rhizosphere microbiome of Beta vulgaris ssp. maritima - ancestor of all beet crops - and modern sugar beets. *Front Microbiol* 5:415.

Zgadzaj R, Garrido-Oter R, Jensen DB, Koprivova A, Schulze-Lefert P, Radutoiu S. 2016. Root nodule symbiosis in *Lotus japonicus* drives the establishment of distinctive rhizosphere, root, and nodule bacterial communities. *Proc Natl Acad Sci U S A* 113:E7996-E8005.

Zhalnina K, Louie KB, Hao Z, Mansoori N, da Rocha UN, Shi S, *et al*. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nat Microbiol* 3(4):470-480.

Zhang N, Wang D, Liu Y, Li S, Shen Q, Zhang R. 2014. Effects of different plant root exudates and their organic acid components on chemotaxis, biofilm formation and colonization by beneficial rhizosphere-associated bacterial strains. *Plant Soil* 374:689-700.

Zhao J, Bodner G, Rewald B, Leitner D, Nagel KA, Nakhforoosh A. 2017. Root architecture simulation improves the inference from seedling root phenotyping towards mature root systems. *J Exp Bot* 68(5):965-982.

Zhu Y, McBride MJ. 2017. The unusual cellulose utilization system of the aerobic soil bacterium *Cytophaga hutchinsonii*. *Appl Microbiol Biotechnol* 101(19):7113-7127.

Zhu Y-G, Smith SE, Barritt AR, Smith FA. 2001. Phosphorus (P) efficiencies and mycorrhizal responsiveness of old and modern wheat cultivars. *Plant Soil* 237:249-255.

Zwanenburg B, Pospíšil T, Ćavar Zeljković S. 2016. Strigolactones: new plant hormones in action. *Planta* 243(6):1311-26.

Summary

Plants rely, in part, on the association with microbes for a number of life-support functions including nutrient acquisition, tolerance to biotic and abiotic stresses and immune regulation. In my thesis, I hypothesized that modern cultivars of crop plants have lost some of the traits needed to recruit host-specific microbiota as compared to their wild relatives, which are genetically more diverse and adapted to pre-agricultural soils. The research presented in this thesis assessed the impact of domestication of common bean (*Phaseolus vulgaris*) and habitat expansion on spermosphere and rhizosphere microbiome composition. For this, the microbiome composition of the spermosphere and the rhizosphere of wild and modern bean accessions grown in an agricultural and a native soil from Colombia, one of the centres of common bean diversification, was characterized by metagenomics and cultivation-dependent approaches.

A higher relative abundance of Bacteroidetes, mainly *Chitinophagaceae* and *Cytophagaceae*, was observed in the rhizosphere of wild bean accessions while an increase in relative abundance of Actinobacteria and Proteobacteria was observed in the rhizosphere of modern bean accessions. These divergences in rhizobacterial community composition between wild and modern bean accessions associated with differences in root morphological traits. In particular the specific root length (SRL) explained a significant portion of the variability in the abundance of families from the phylum Bacteroidetes. Using 16S-rDNA data from several other studies revealed that also wild relatives of other crop plant species presented higher relative abundances of members of the Bacteroidetes. The reason behind the higher relative abundance of Bacteroidetes in the root and rhizosphere compartments of wild relatives of various plant species is yet unknown but may be related to their ability to metabolize more complex polymers in the root exudates.

Plant domestication not only comes with changes in plant traits, but is also accompanied by progressive changes in the habitat and crop management practices. Hence, the transition from native habitats to agricultural soils may have led to a loss of plant-associated microbes thereby affecting specific, co-evolved beneficial functions of the plant microbiome. To begin to understand how these domestication factors affect the rhizosphere microbiome, the bean accessions were grown in a native and an agricultural soil from the Colombian highlands. The results showed that this transition of common bean from a native soil to an agricultural soil led to a gain of rhizobacterial diversity and led to a stronger bean genotype-dependent effect on rhizosphere microbiome assembly. A core rhizosphere microbiome was identified for common bean which mainly consisted of bacterial genera with nitrogen fixing capabilities, an important feature of microbes associated with leguminous plant species. However, also for other non-leguminous plant species, these bacterial genera are members of the core rhizosphere microbiome, suggesting a homogenization of rhizobacterial diversity of plants grown in different agricultural landscapes. Co-occurrence analyses further showed a reduction in complexity of the interactions going from native to agricultural soil which may indicate that rhizobacterial community assembly for common bean grown in agricultural soil is more modular than for common bean in native soil, making it relatively more easy for bacteria and other microorganisms to invade and establish in the rhizosphere of bean plants grown in agricultural soils.

The results of this thesis also showed that already during seed imbibition and germination, small but significant differences were detected between the spermosphere microbiomes of wild and modern bean accessions. These results suggest a domestication effect on microbiome assembly already at this early developmental stage. Exudation profiles showed that the spermosphere of a modern bean accession contained higher levels of

glutamate and glutamine as compared to a wild bean accession. Although a strong impact on microbiome composition was observed, a clear trend towards specific bacterial taxa consistently enriched in vitro by this amino acid was not observed. These results suggested that more complex mixtures of different exudate constituents are needed to mimic in vitro the selection occurring in situ in the bean spermosphere.

In conclusion, the research presented in this thesis showed that domestication of common bean had a significant effect on the composition of the rhizosphere and spermosphere microbiome. The higher abundance of Bacteroidetes in the rhizosphere of wild relatives of common bean and other crop species as well as the higher abundance of Actinobacterial families in the rhizosphere of modern crop cultivars provide new research directions to unravel the chemical cross-talk between plants and microbes, and to investigate the functional impact of these microbiome shifts on plant growth, development and health.

Samenvatting

Planten zijn deels afhankelijk van microorganismen voor een aantal functies waaronder de acquisitie van nutriënten, de bescherming tegen abiotische en biotische stress en de regulatie van de immuunrespons. De hypothese van mijn proefschrift is dat moderne cultivars van gewassen als gevolg van de domesticatie een aantal eigenschappen hebben verloren die nodig zijn voor het recruteren van specifieke microorganismen in vergelijking met hun wilde verwanten die genetisch meer divers zijn en beter aangepast aan natieve gronden. In het onderzoek beschreven in dit proefschrift is onderzocht wat de impact is van domesticatie van boon (*Phaseolus vulgaris*) op de samenstelling van het microbioom van de spermosfeer en rhizosfeer. Daartoe werd met behulp van metagenomics en klassieke technieken de microbioomsamenstelling bepaald van wilde en moderne accessies van boon in landbouwgrond en natieve gronden van Colombia, een van de diversificatie centra van boon.

Voor de rhizosfeer van wilde bonen werd een hogere relatieve abundantie van de Bacteroidetes waargenomen, met name van *Chitinophagaceae,* terwijl Actinobacteria en Proteobacteria meer voorkwamen in de rhizosfeer van moderne cultivars van boon. Deze verschillen in de samenstelling van de bacteriële gemeenschappen in de rhizosfeer waren geassocieerd met verschillen in morfologische eigenschappen van het wortelstelsel. Met name de specifieke wortellengte (SRL) verklaarde een significant deel van de variatie in abundantie van bacteriële Bacteroidetes families. Op basis van de 16S-rDNA data van diverse andere studies waren de Bacteroidetes ook in een hogere abundantie aanwezig in de rhizosfeer van wilde verwanten van andere plantensoorten. Het onderliggende mechanisme is nog niet bekend maar zou gerelateerd kunnen zijn aan het vermogen van Bacteroidetes om complexe polymeren in wortelexudaten van wilde verwanten te metaboliseren.

Domesticatie van planten gaat niet alleen gepaard met veranderingen in planteneigenschappen maar ook met veranderingen in habitat en teeltmaatregelen. Daarom kan de transitie van de natieve habitat naar landbouwgronden mogelijk geleid hebben tot een verlies van plant-geassocieerde microorganismen en specifieke goedaardige functies van het plantenmicrobioom. Om te begrijpen hoe deze domesticatiefactoren het rhizosfeermicrobioom hebben beïnvloed werden de boon accessies opgekweekt in natieve en landbouwgronden van de Colombiaanse hooglanden. De resultaten lieten zien dat de transitie van natieve gronden naar landbouwgrond heeft geleid tot een toename in bacteriële diversiteit en tot een sterker genotype-afhankelijke samenstelling van het rhizosfeermicrobioom. Het gemeenschappelijke rhizosfeermicrobioom van boon bestond voornamelijk uit bacteriële genera met het vermogen om stikstof te fixeren, een belangrijke eigenschap van microorganismen geassocieerd met vlinderbloemige plantensoorten. Echter, ook voor niet-vlinderbloemige soorten waren deze bacteriële genera onderdeel van het gemeenschappelijk microbioom, hetgeen wijst op een homogenisatie van de diversiteit van bacteriële gemeenschappen geassocieerd met de rhizosfeer van planten in verschillende landbouwgronden. Netwerk analyses lieten een reductie in complexiteit van interacties zien tussen natieve en landbouwgronden, hetgeen kan betekenen dat de samenstelling van de bacteriële gemeenschap in landbouwgrond meer modulair is dan in natieve grond. Modulariteit kan het gemakkelijker maken voor bacteriën en andere microorganismen om de rhizosfeer binnen te dringen en zich te vestigen.

De resultaten van dit proefschrift toonden tevens aan dat al tijdens de imbibitie en kieming van bonenzaden kleine doch significante verschillen werden waargenomen tussen het spermosfeermicrobioom van wilde en moderne accessies. Deze resultaten suggereren dat

het effect van domesticatie al plaatsvindt tijdens deze vroege ontwikkelingsfase. Exudatie profielen lieten zien dat de spermosfeer van moderne bonen hogere concentraties bevatten van glutamaat en glutamine in vergelijking met de wilde bonen. Desondanks vonden we niet een duidelijk effect van deze aminozuren op de microbioomsamenstelling. Waarschijnlijk is een meer complexe mix van verschillende stoffen in de exudaten nodig voor de selectie die plaatsvindt in de spermosfeer van boon.

Concluderend toonde het onderzoek in dit proefschrift aan dat domesticatie van boon een significant effect heeft gehad op de samenstelling van het rhizosfeer- en spermosfeermicrobioom. De hogere dichtheid van Bacteroidetes in de rhizosfeer van wilde verwanten van boon en andere plantensoorten en de hogere dichtheid van Actinobacteria in de rhizosfeer van moderne cultivars van boon geven nieuwe richtingen in het onderzoek naar de chemische communicatie tussen planten en microben en naar de functionele impact van deze microbioomverschuivingen op de groei, ontwikkeling en gezondheid van planten.

Acknowledgements

I feel deeply grateful with the Colombian estate and with those who implemented the government program of PhD fellowships that has been maintained with public funding during the last decade. Thanks to this program I had the opportunity to live, study and perform a research project in the Netherlands.

My gratitude to all the people and institutions that were involved in the elaboration of the different chapters presented in this thesis. In Brazil, thanks to Embrapa Meio Ambiente and to Instituto Agronômico de Campinas. Special thanks to Rodrigo, which had a great influence in early stages of the project. Maike, Harold, Josiane, Stalin, Carol and to all the people at embrapa and IAC that helped me during my stay in Brazil, many thanks. Also in Brazil, thanks to Felipe and Antonio from the Escola Superior de Agricultura Luiz de Queiroz from the Universidade de São Paulo. In Colombia, thanks to Camilo Ramírez and the BA&A lab from Universidad de Antioquia, and Daniel Debouck and Orlando Toro at the International Center for Tropical Agriculture. Also in Colombia thanks to my brothers, Hugo and Jorge, for helping me with soil collection, and to Tavo for traveling with our team across Colombia. I would like to thank also the people at Phytopathology Lab at Wageningen University, in which I was initially appointed as PhD student. Thanks a lot to Ester for her help with the logistics of the initial experiments. At NIOO I am very grateful to Mattias that with his knowledge and skills became a key person in the elaboration of this thesis. Thanks also to Ciska for her great support during the chemical determinations. Thanks also to Gregor and Roel for their support during greenhouse experiments and to the secretary department for their valuable help with administrative stuff.

I also would like to specially thank Desalegn, Mauro and Thiago, great friends that thanks to NIOO I had the opportunity to meet, and that shared their time and thoughts (about science, but mostly about many other subjects) with me during all these years in the Netherlands. Similarly, I would like to express my sincere gratitude to Ruth, Natalia, Viviane, Lucas, Je-Seung, Kristin, Olaf, Nurmi, Xu, Chunxu, Irene, Adam and Kay. I tried to learn from the strengths of each one of you, very talented and hard-working young scientists with whom I had the opportunity to work and share during my PhD. Extensive gratitude to people of Jos group and ME in general, with whom I had the opportunity to share meetings, coffee breaks and borrels during the last 5 years.

My gratitude also for the Colombian community in Wageningen: Jose, Amalia, Leo, Daniel, Diana, Andrea, Juan, Cesar and their families. You were a constant connection with Colombia, and your friendship was a relief in tough moments. I would like to thank also to GVC, the football club I belonged during my whole stay in the Netherlands. Trainings and matches were more than time practicing a sport, and instead became a way to connect with Dutch culture. Imre, Ruben, Joost, Francesco, Douwe, and all those players with whom I had the opportunity to play, thanks a lot for those glorious (less frequent) and tough (quite frequent) moments in the pitch.

I would like to dedicate this thesis to Silvia and her family. Their company, help and support was invaluable. Thanks for all my years living in Europe, you showed me what a family should be. Silvia, this thesis is especially dedicated to you, because only you really knew how important this was for me. I will be always grateful. Special thanks to Iván, great friend and great support. I learned a lot from you about the finest details in life. Finally, I would like to thank my mother and my two brothers. You have been always there, and surely will be always there.


WILL/
MAUER
MER !
1990
VAUDE

About the author

Juan Esteban Pérez Jaramillo was born on 24th of March 1984 in Jericó, province of Antioquia, Colombia. In 2008 he concluded his undergraduate studies in Biology at the University of Antioquia. Then he started a Master in Biology at the same university, where he studied the vectors and reservoirs of leishmaniasis in the Darien region of Colombia. In 2011, he was awarded with an International Talent fellowship from the Navarra's Government in Spain to study a Master in Agrobiology in the Public University of Navarra. During his Master studies he worked under the supervision of Prof. Dr. Pedro Aparicio Tejo, and studied the effect of organic matter amendments, more specifically treated sewage sludges, on the enzymatic activities of the soil as well as in the N_2O emissions. After the conclusion of his studies in Spain, he went back to Colombia for a short period of time, where he was awarded with a PhD fellowship from the Colombian Department of Science, Technology and Innovation to study in an international university. Consequently, in 2013 he travelled to the Netherlands and started his doctoral research under the supervision of Prof. Dr. Jos Raaijmakers, initially at the Laboratory of Phytopathology of Wageningen University, and from 2014 at the Netherlands Institute of Ecology. The findings of his research are presented in this thesis. Early in 2018, he went back to Colombia where he started a position as research associate at the Program for the Study and Control of Tropical Diseases (PECET) from the University of Antioquia. Since February of 2019 he started a position as Professor of Biotechnology in the University of Antioquia, where he develops activities of teaching and research.

List of Publications

Pérez-Jaramillo JE, de Hollander M, Ramírez CA, Mendes R, Raaijmakers JM, Carrión VJ. Deciphering the microbiome assembly of wild and modern common bean (*Phaseolus vulgaris*) grown in native and agricultural soils from Colombia. (*Chapter 4*, *submitted*)

Carrión VJ, **Pérez-Jaramillo JE**, Cordovez V, de Hollander M, Tracanna V, Mendes LW, *et al*. Activation of the endophytic microbiome by fungal pathogen invasion. (*submitted*)

Rossmann M, **Pérez-Jaramillo JE**, Chiaramonte JB, Kavamura VN, Dumack K, Fiore-Donno AM, *et al*. Cercozoa as a key factor in rhizosphere microbiome assembly. (*submitted*)

Hannula SE, Ma H, **Pérez-Jaramillo JE**, Pineda A, Bezemer M. Structure and ecological function of the soil microbiome affecting plant-soil feedbacks in the presence of a soil-borne pathogen. (*to be submitted*)

Pérez-Jaramillo JE, Carrión VJ, de Hollander M, Raaijmakers JM. 2018. A walk on the wild side of plant microbiomes. Microbiome. 6:143.

Pérez-Jaramillo JE, Carrión VJ, Bosse M, Ventorim Ferrão LF, de Hollander M, Franco Garcia AA, et al. 2017. Linking rhizosphere microbiome composition of wild and domesticated *Phaseolus vulgaris* to genotypic and root phenotypic traits. ISME Journal 11, 2244–2257.

Pérez-Jaramillo JE, Mendes R, Raaijmakers JM. 2016. Impact of plant domestication on rhizosphere microbiome assembly and functions. Plant Molecular Biology 90: 635–644.

Calleja-Cervantes ME, Irigoyen I, Gorriz C, **Pérez-Jaramillo JE**, Irañeta J, Amorena A, Aparicio-Tejo PM, Menéndez S. 2013. Twenty years of continued application of treated sewage sludge: nitrous oxide emissions induced in agricultural soils. Conference paper. Ramiran. p.S2, 29. ISBN: 978-2-7380-1337-8

Carrillo LM, Betancur S, Roldán D, **Pérez-Jaramillo JE**, Galeano D, Loaiza ET, Giraldo CA. 2012. Implementation of a PCR-based method for the diagnosis of Ehrlichia spp, in canine in Medellin (Colombia). Revista CES Medicina Veterinaria y Zootecnia 7(2):38-46.

Vivero RJ, Bejarano EE, Castro M, Vélez A, **Pérez-Jaramillo JE**, Pérez-Doria A, Vélez ID, Carrillo LM. 2010. Thirteen new records of *Lutzomyia* (Diptera: Psychodidae) for the department of Vichada, Colombian Orinoquia. Biota Neotropica 10(2): 401-404.

The research described in this thesis was performed at the Department of Microbial Ecology of the Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, The Netherlands; at Universidad de Antioquia, Medellín, Colombia, and the Brazilian Agriculture Research Corporation, Embrapa Meio Ambiente, Jaguariúna, São Paulo State, Brazil. The doctoral research was financially supported by the Colombian Department of Science, Technology and Innovation – COLCIENCIAS, through the doctoral grant 568-2012-15517825.

This is NIOO-thesis number 166